

The Jet Image



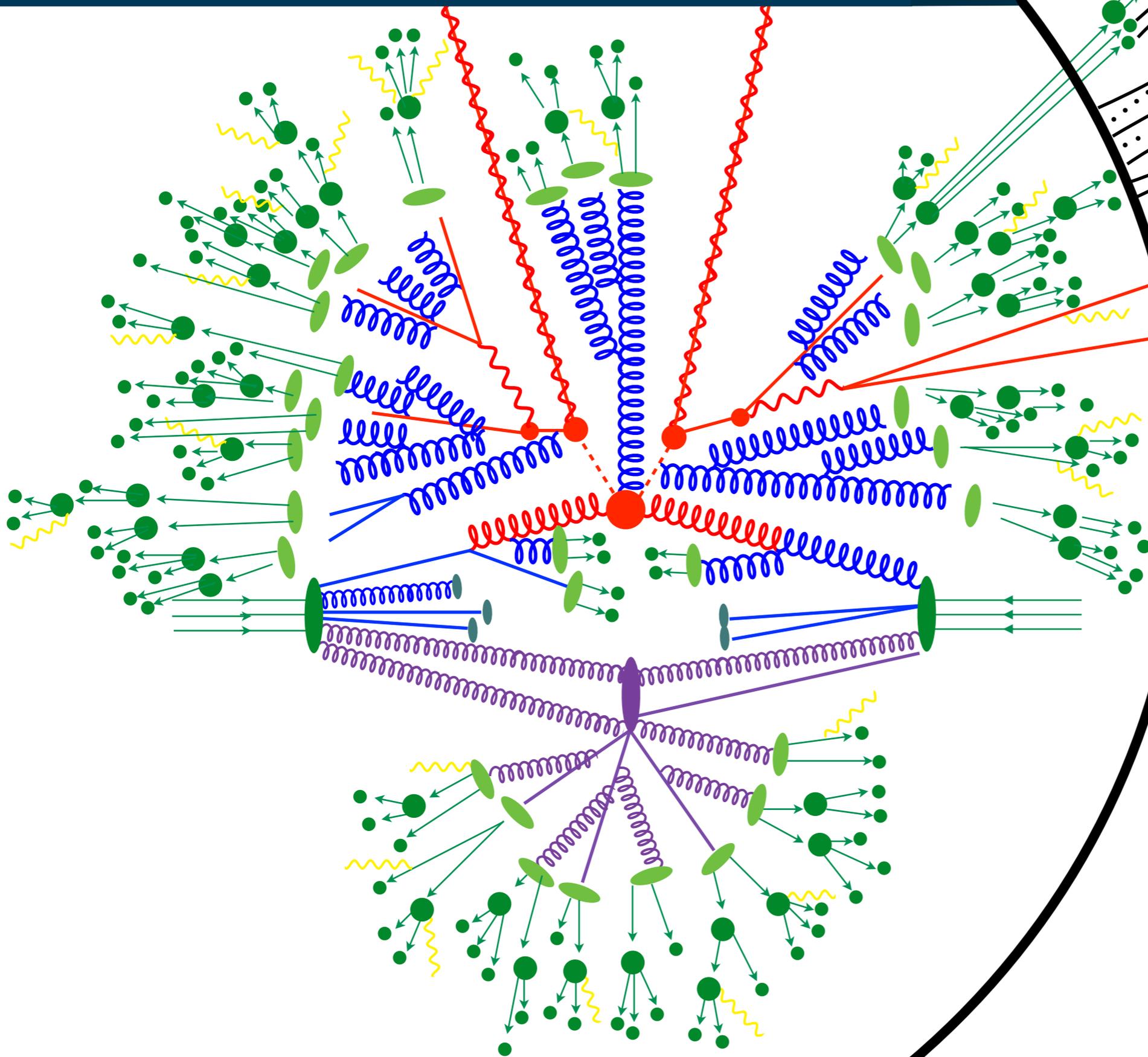
Benjamin Nachman

Lawrence Berkeley National Laboratory

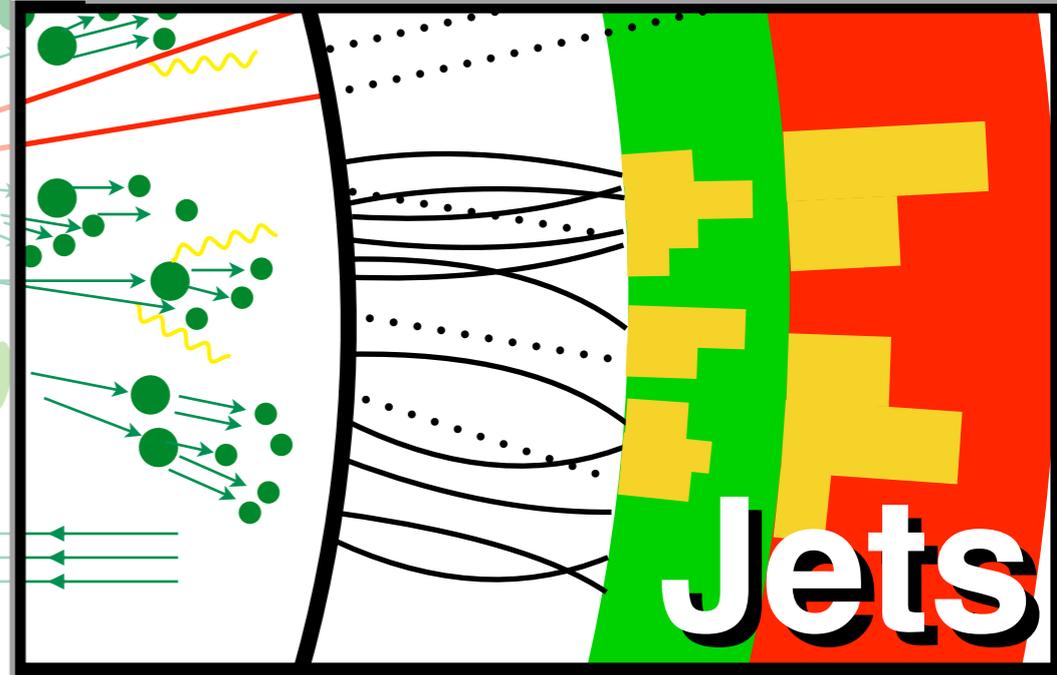
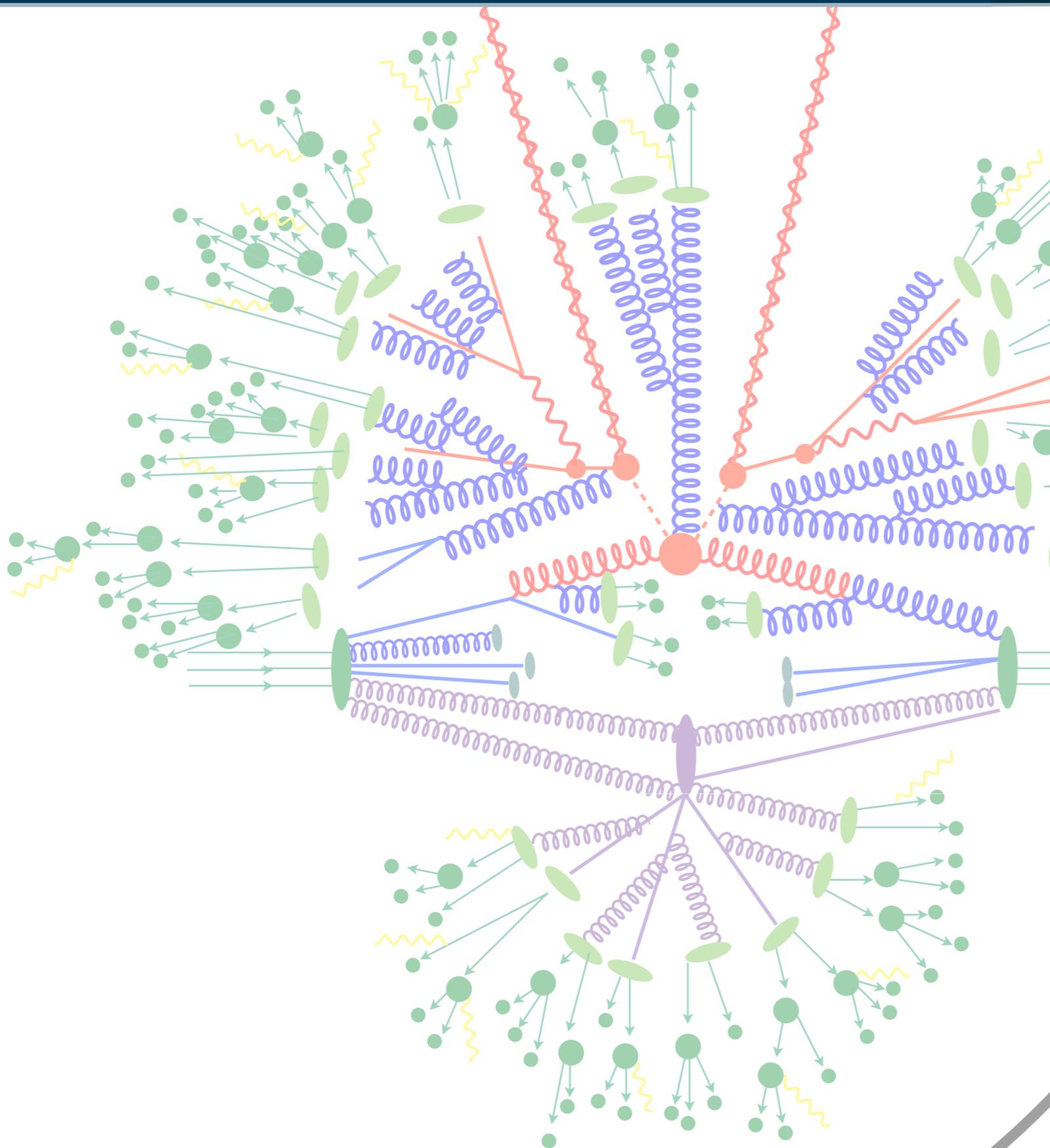
DS@HEP 2017

Outline: Pre-processing \rightsquigarrow (D)NN Classifiers
 \rightsquigarrow Generation \rightsquigarrow Modeling \rightsquigarrow Related work

Quantum Chromodynamics (QCD)

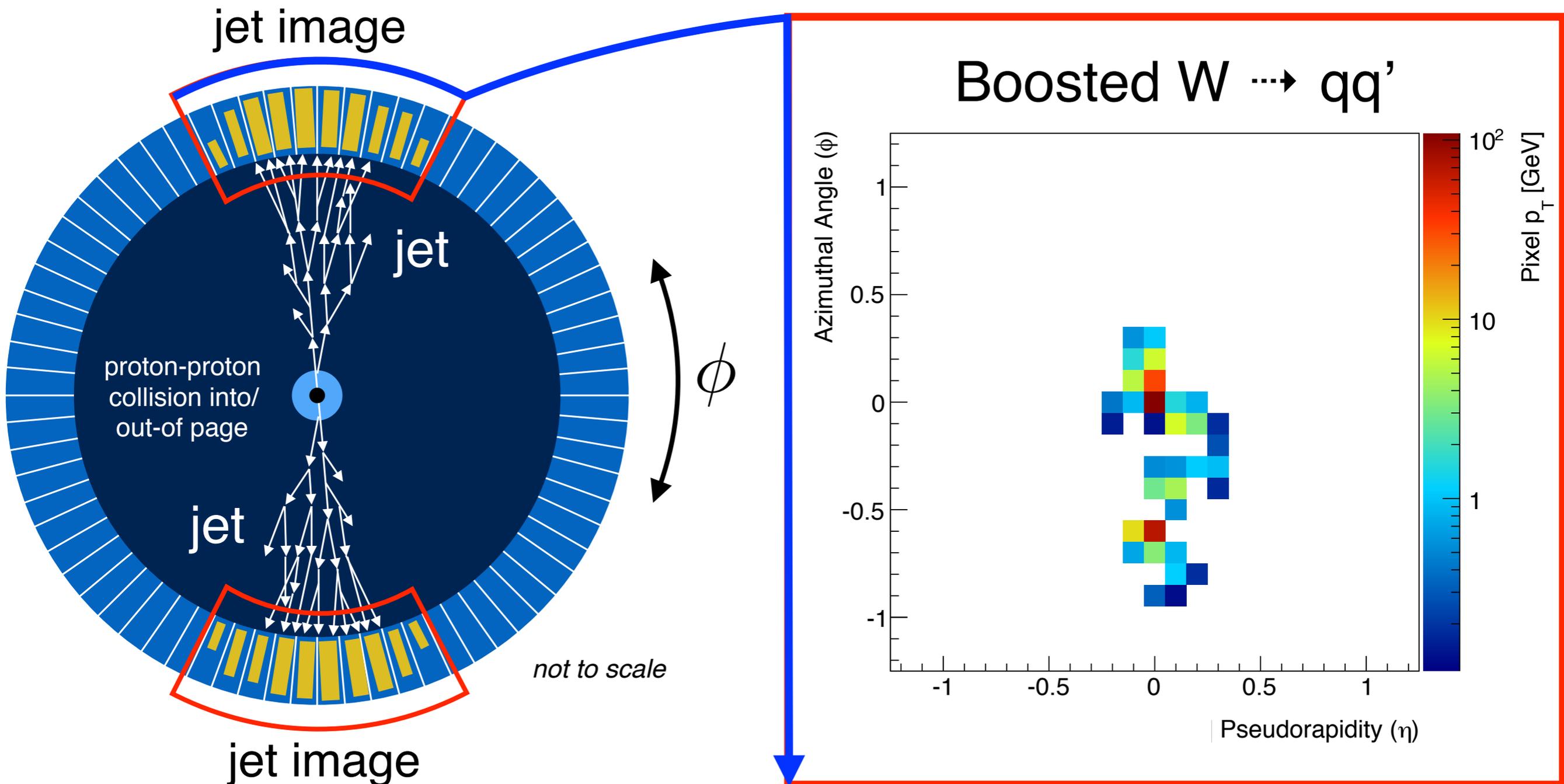


Quantum Chromodynamics (QCD)

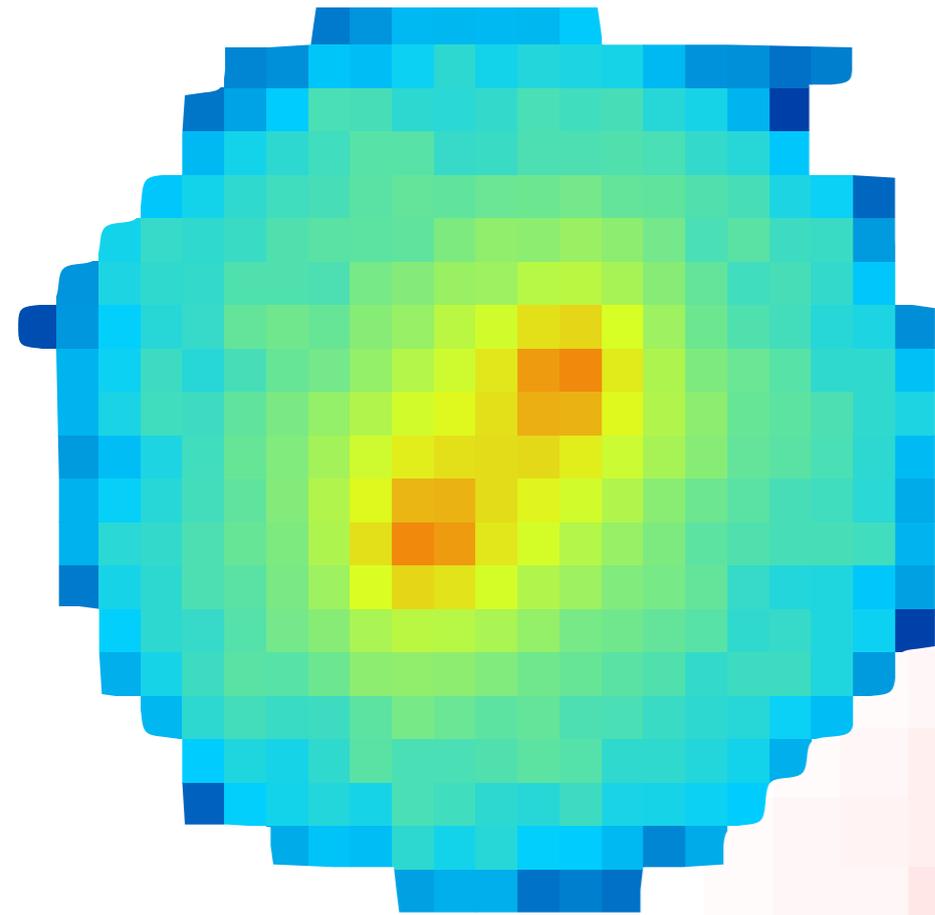


Jets

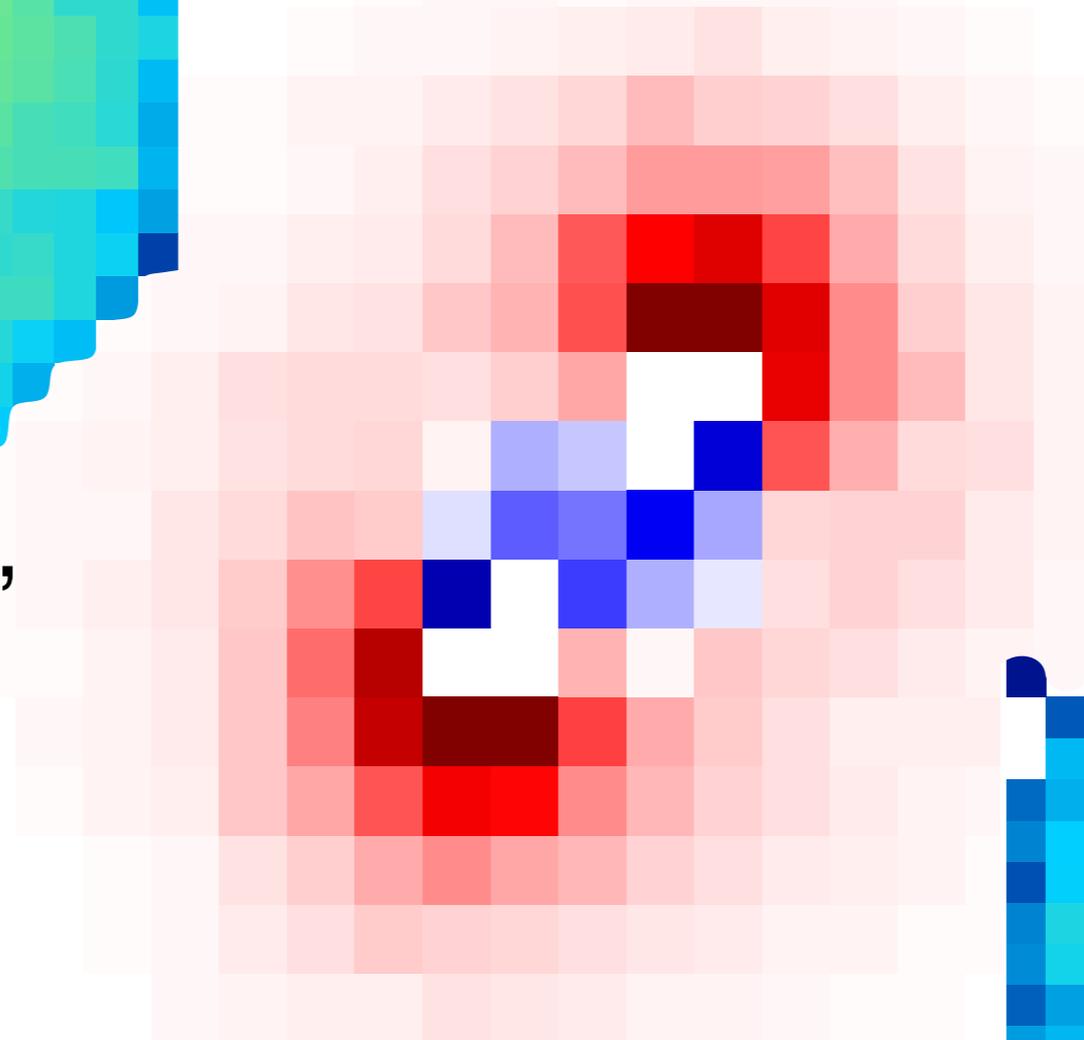
Jet Image: *A two-dimensional fixed representation of the radiation pattern inside a jet*



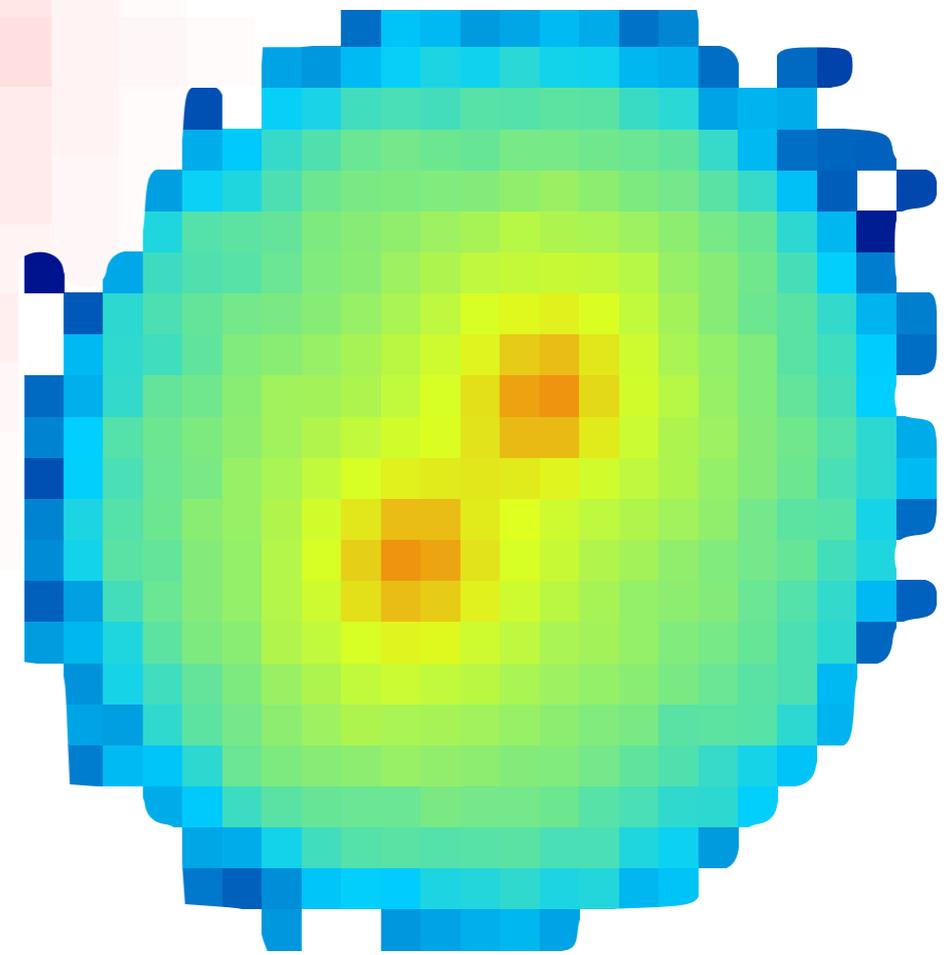
Can directly visualize physics
and we can benefit from the
extensive image processing literature



singlet $\rightsquigarrow qq'$



octet $\rightsquigarrow qq'$



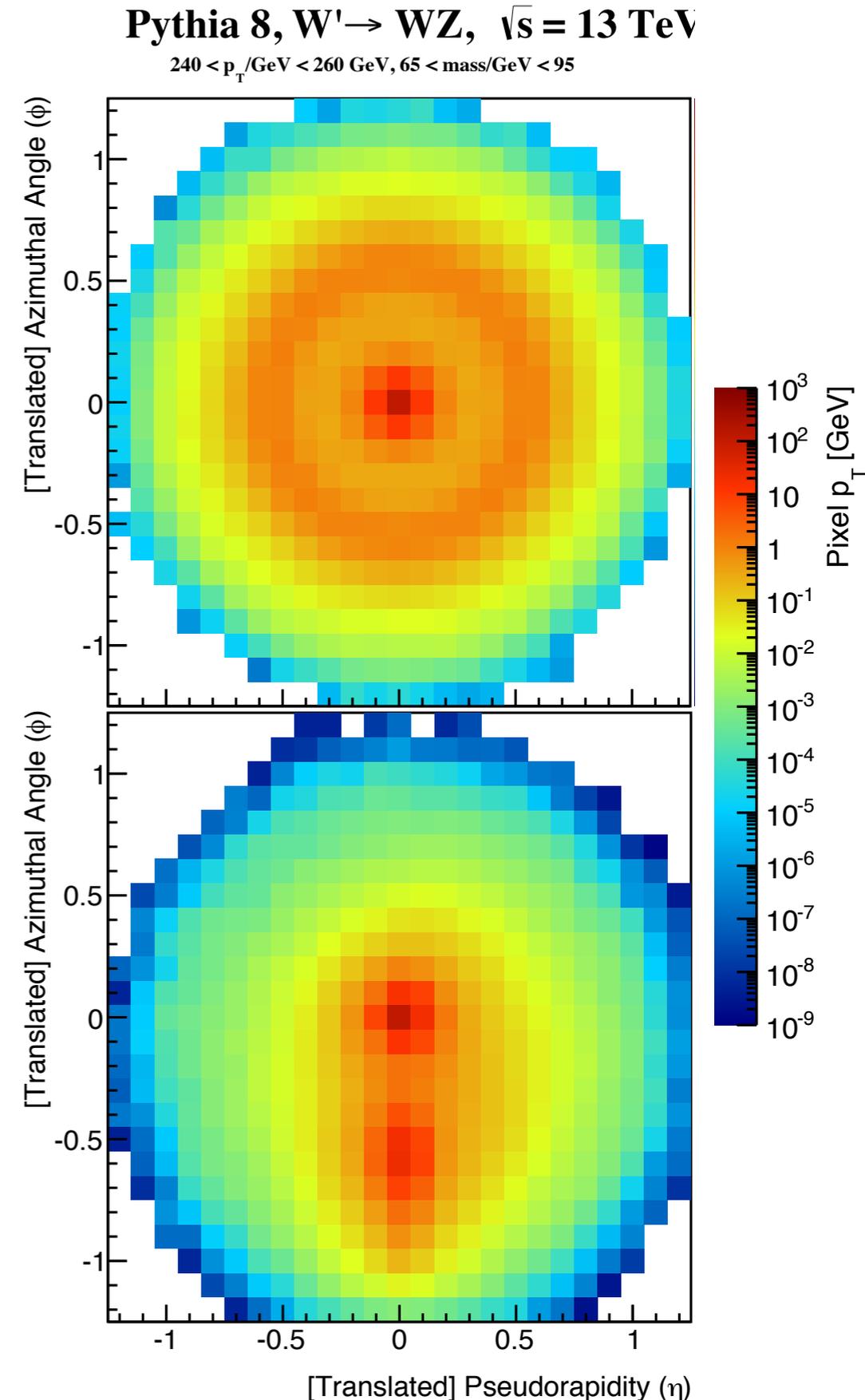
there is information encoded in the
physical distance between pixels
(will mention other fixed representations later)

Pre-processing is an important aspect of image recognition

However, some steps can *damage the physics information content* of a jet image

I'll briefly illustrate some of these challenges in the next slides

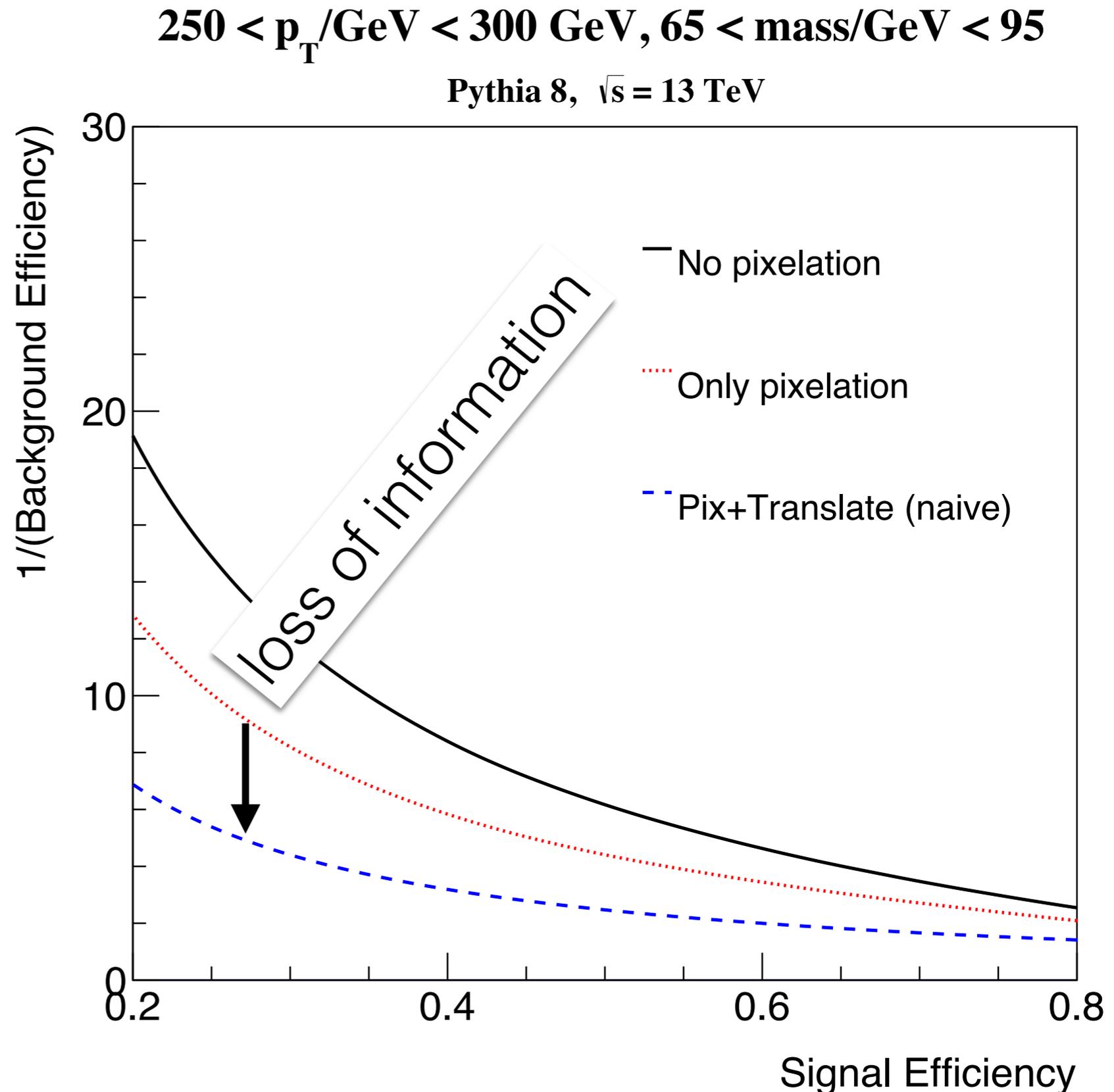
there are also non-standard ideas such as "Zooming" - see *J. Barnard et al. 1609.00607*



For calorimeter images, it is natural to think about energy as intensity.

However, **centering the image** in η corresponds to **boosting along z!**

Therefore, it is very important to use **p_T and not E.**



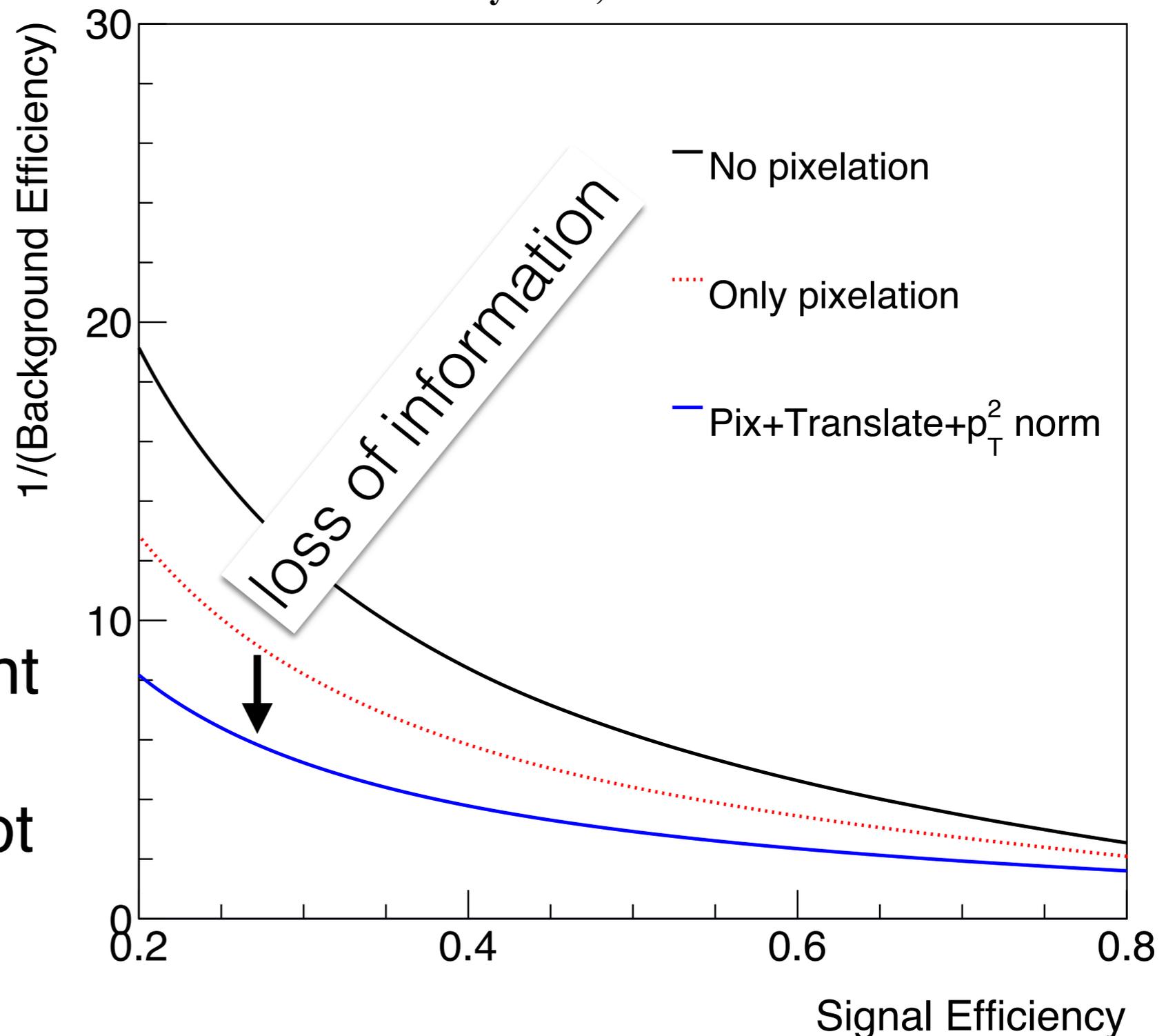
Similar story for image normalization.

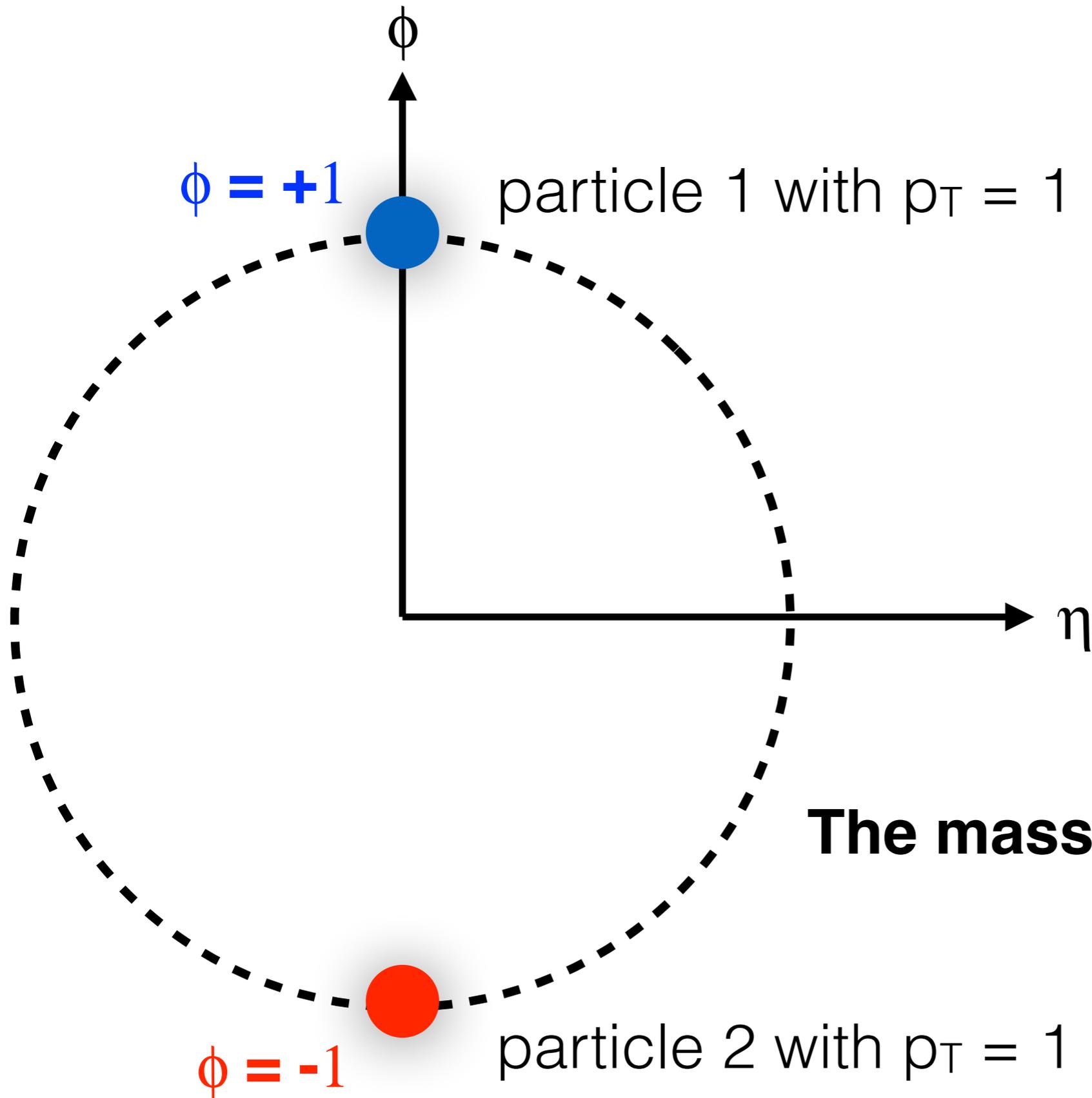
As with E instead of p_T , this adds in \sim random noise to e.g. the mass

Therefore, it is important to do **ensemble-normalizations** and not **jet-by-jet norms**.

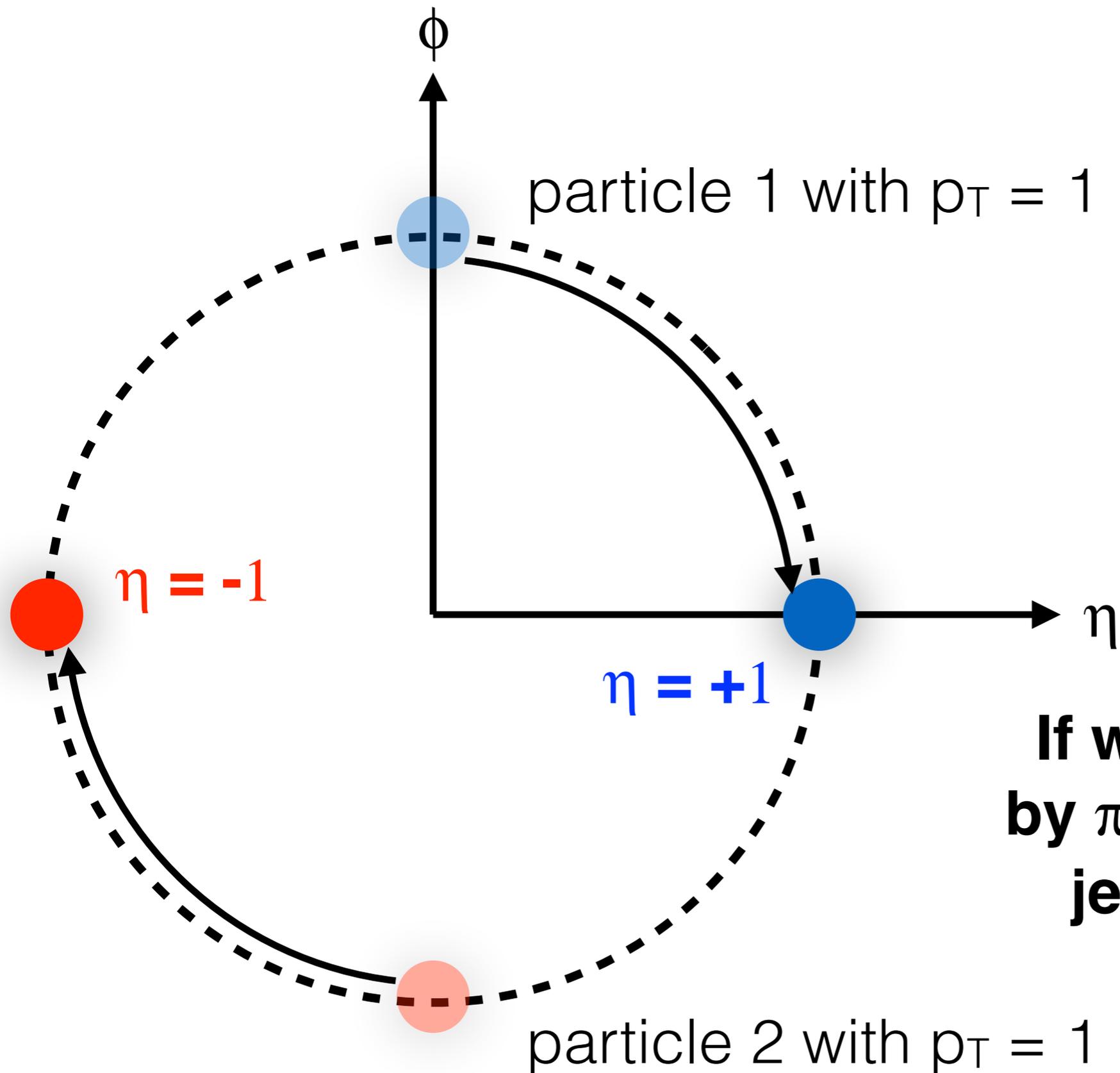
$250 < p_T/\text{GeV} < 300 \text{ GeV}$, $65 < \text{mass}/\text{GeV} < 95$

Pythia 8, $\sqrt{s} = 13 \text{ TeV}$

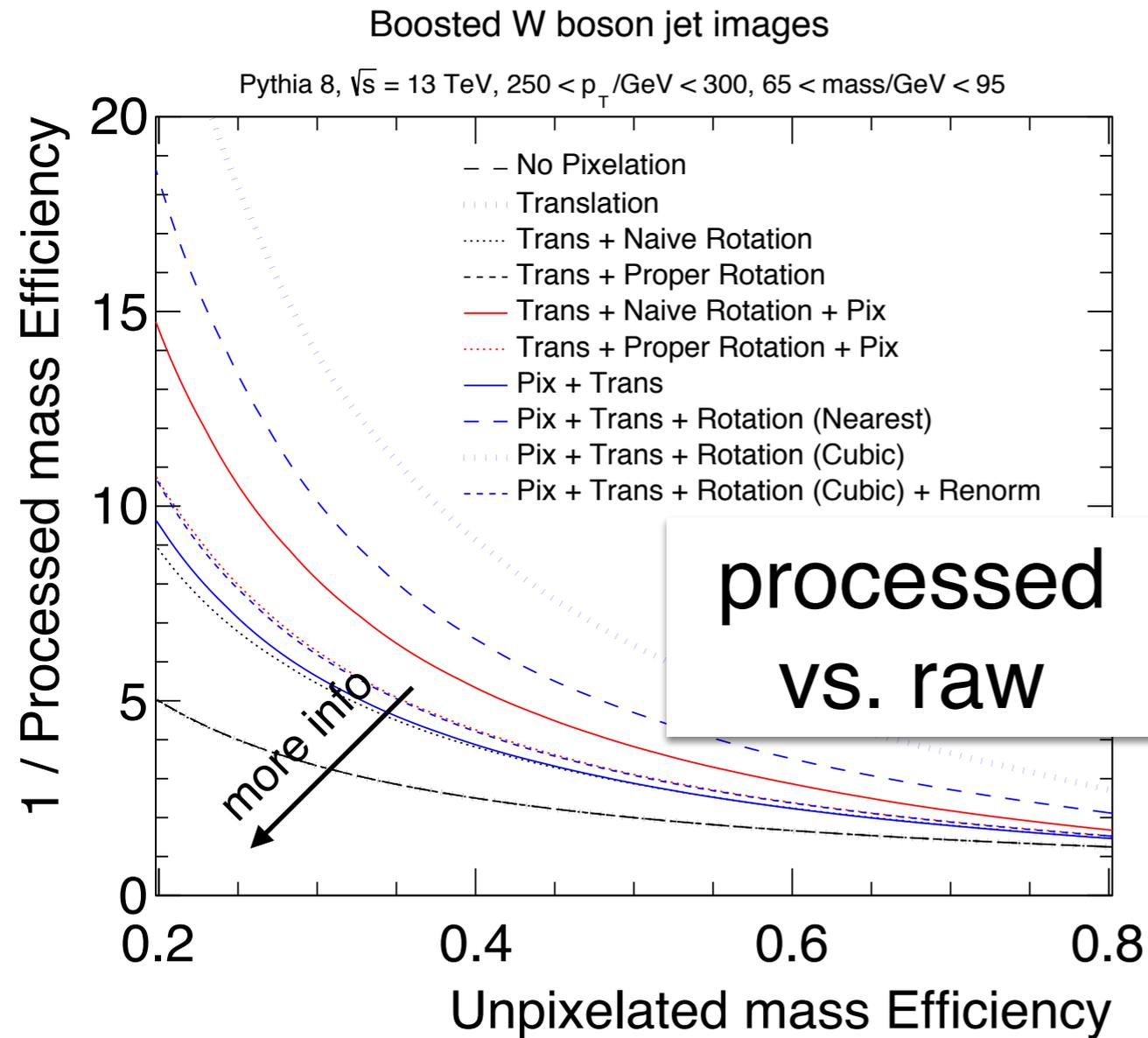




The mass of this 'jet' is ~ 1.7



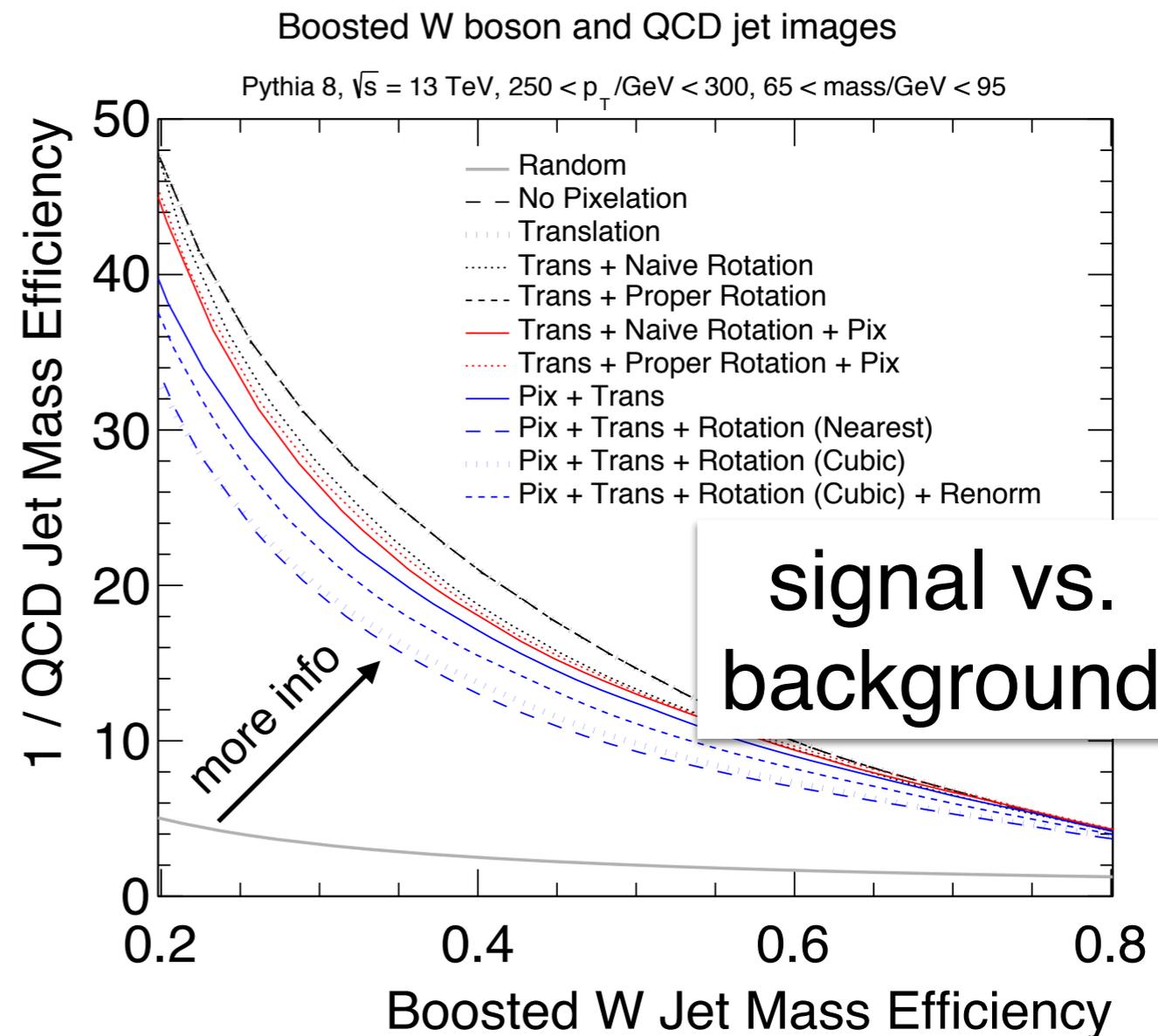
**If we rotate the jet
by $\pi/2$, then the new
jet mass is ~ 2.4**



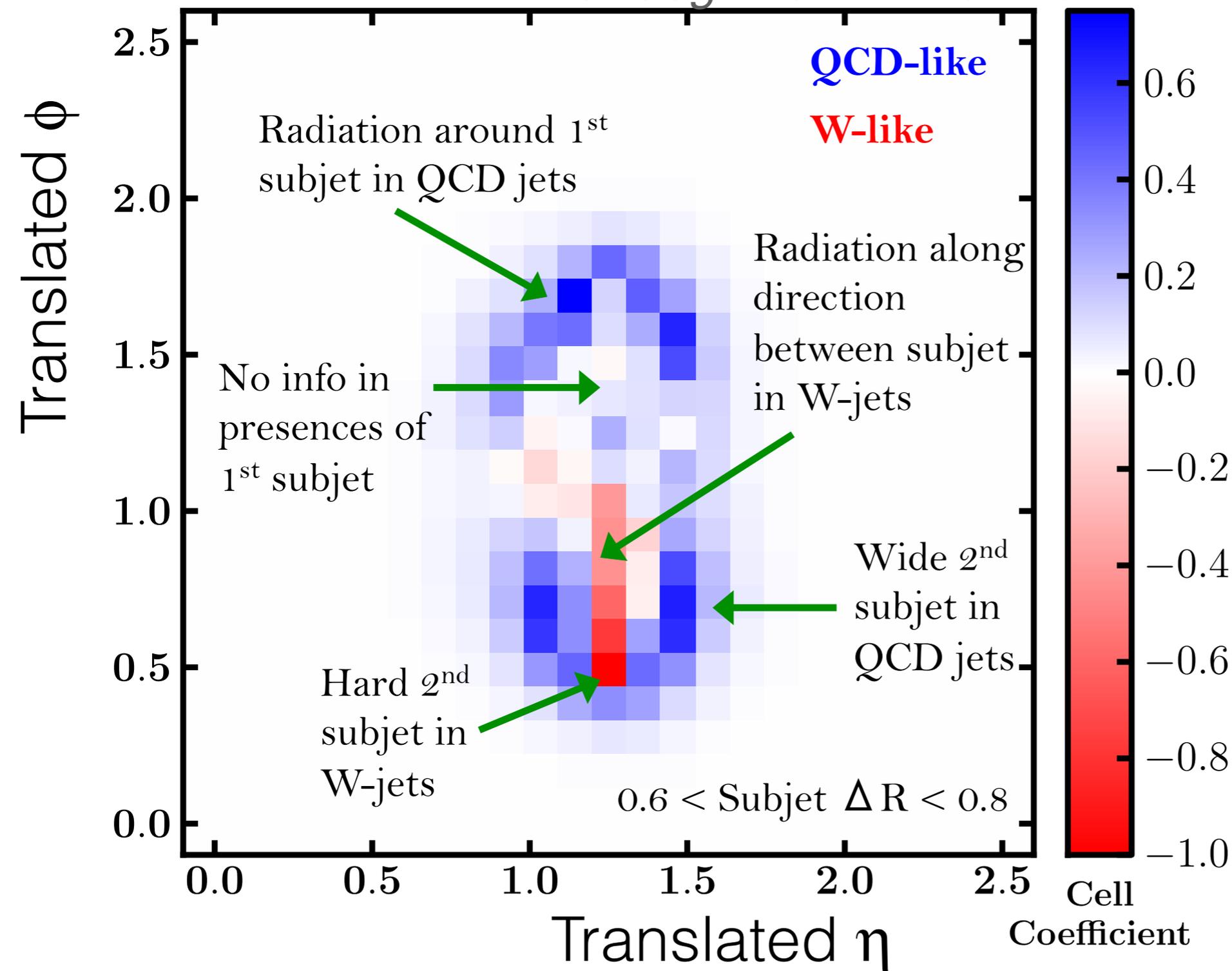
Can add in information to “undo” rotation or augment the dataset with “ghost images” (1612.01551)

de Oliveira et al. arXiv:1701.05927

Can do a ‘proper’ rotation that preserves mass, but changes e.g. τ_{21} .



J. Cogan et al. 1407.5675



“Fisher Jet”

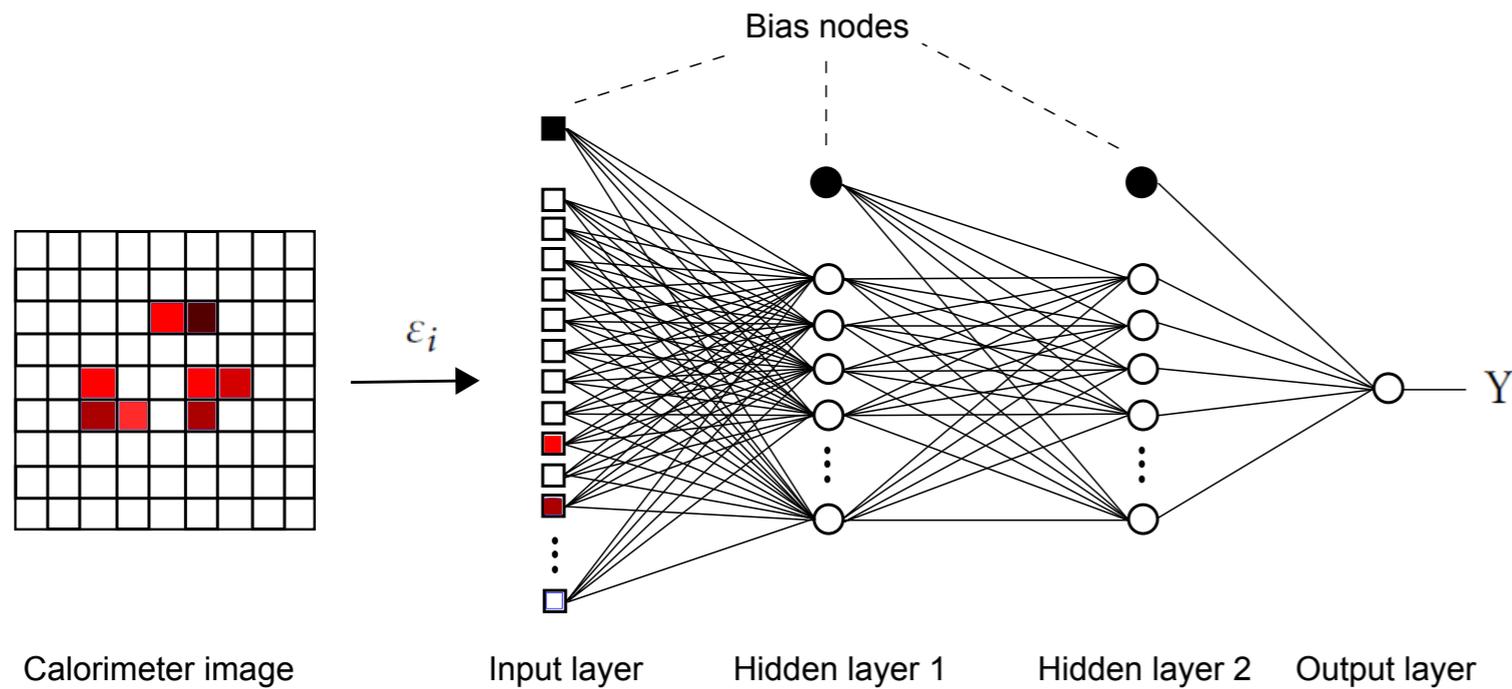
Build discriminant from projecting onto this image

Directly interpretable!

Add in a small non-linearity by binning in ΔR

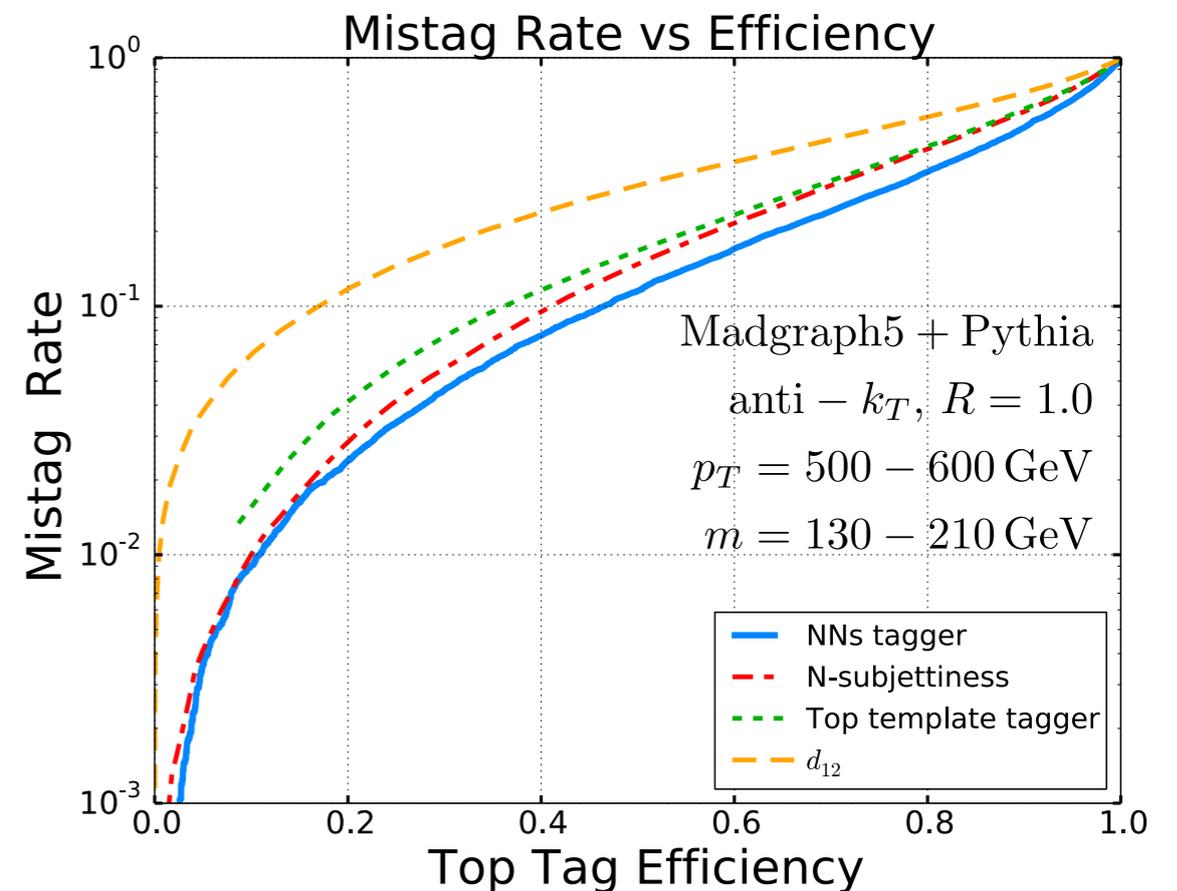
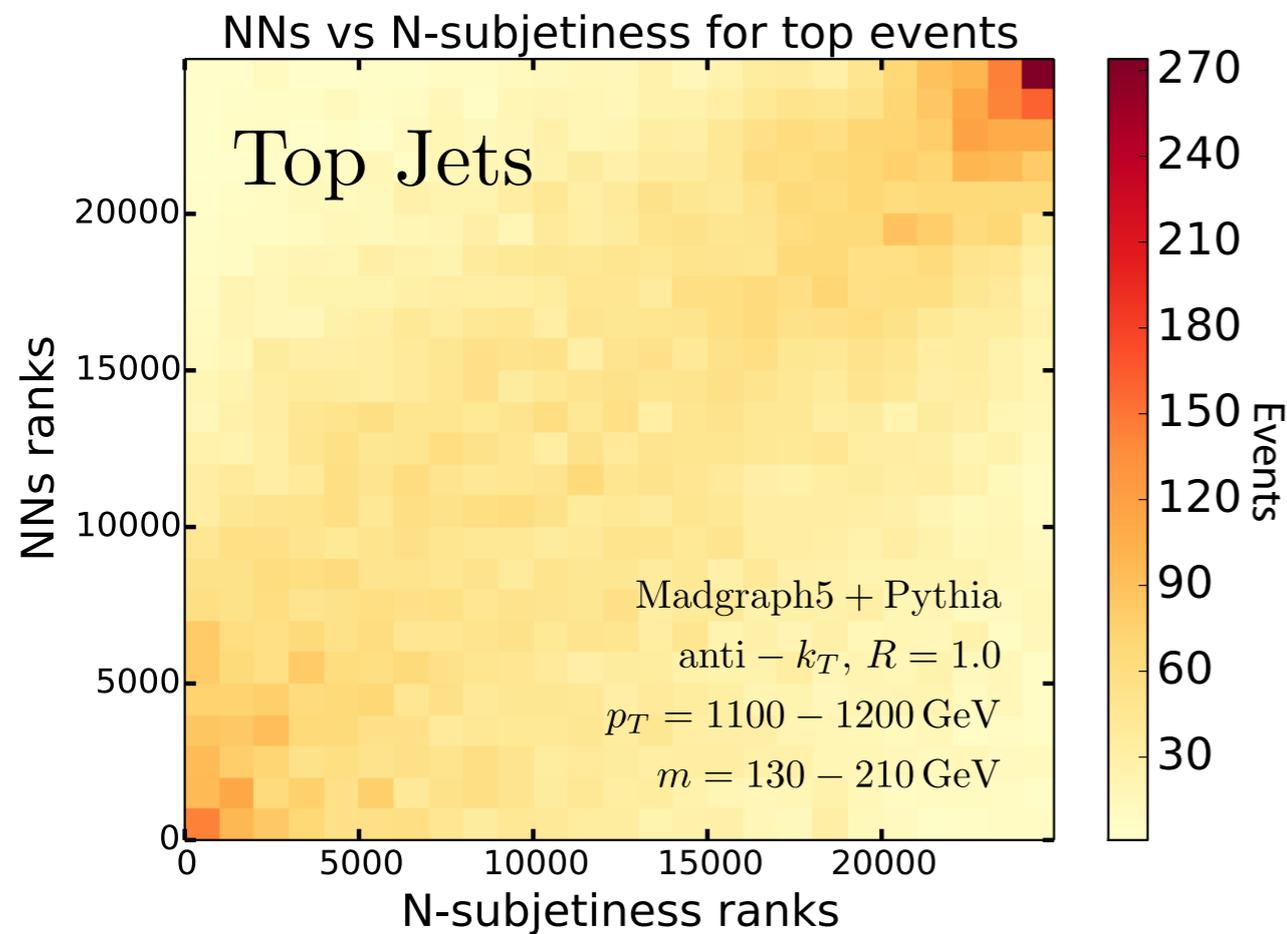
(eyes on a face closer when further away!)

Maximize between class versus within class variance



L. Almeida et al. 1501.05968

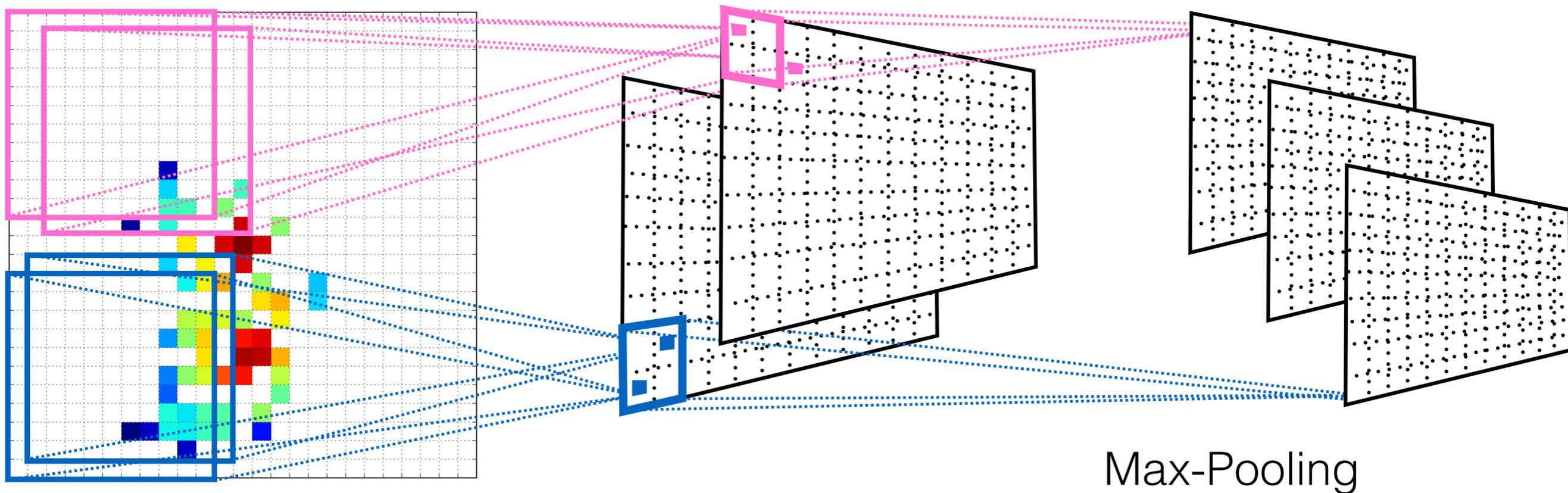
First application of the jet images idea using (shallow) NN's with top-quark tagging



de Oliveira et al. 1511.05190

Convolutions

Convolved Feature Layers



$W' \rightarrow WZ$ event

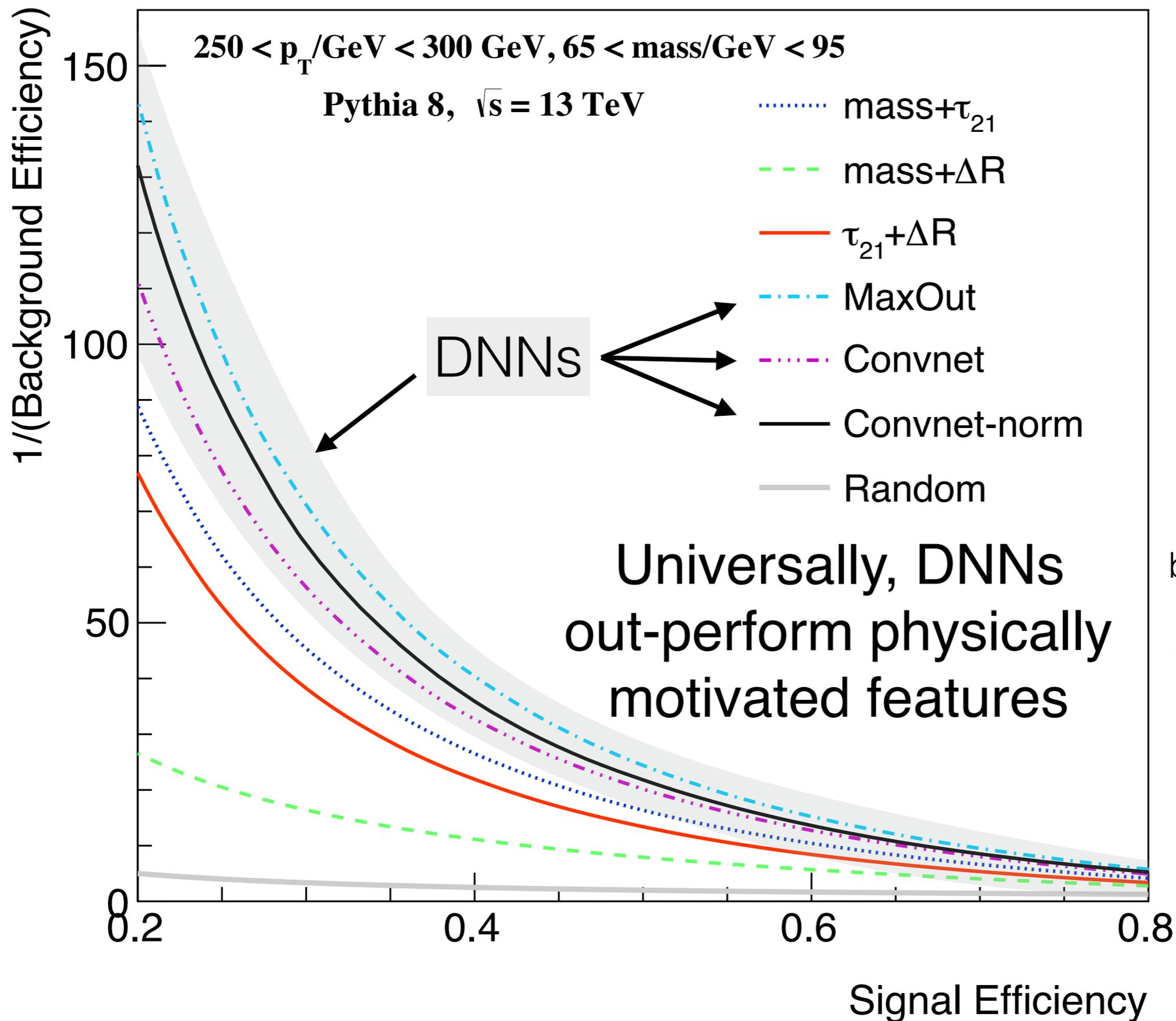
**Subsequent
developments:**

P. Baldi et al. 1603.09349 (W-tagging)

J. Barnard et al. 1609.00607 (W-tagging)

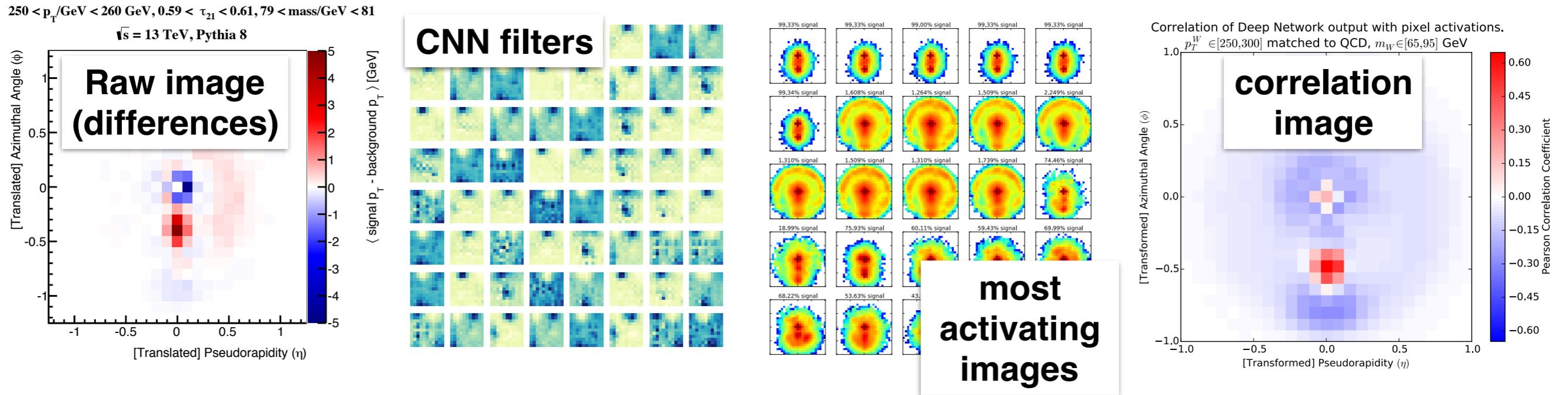
P. Komiske et al. 1612.01551 (q/g-tagging)

G. Kasieczka et al. 1701.08784 (top-tagging)

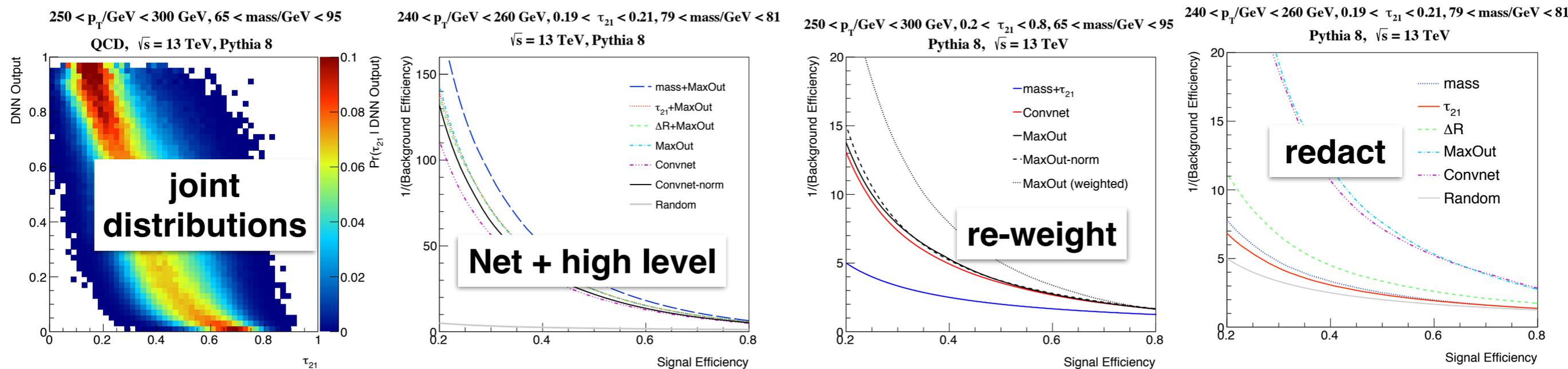


by how much varies by study; also no two groups have the same setup or metrics...

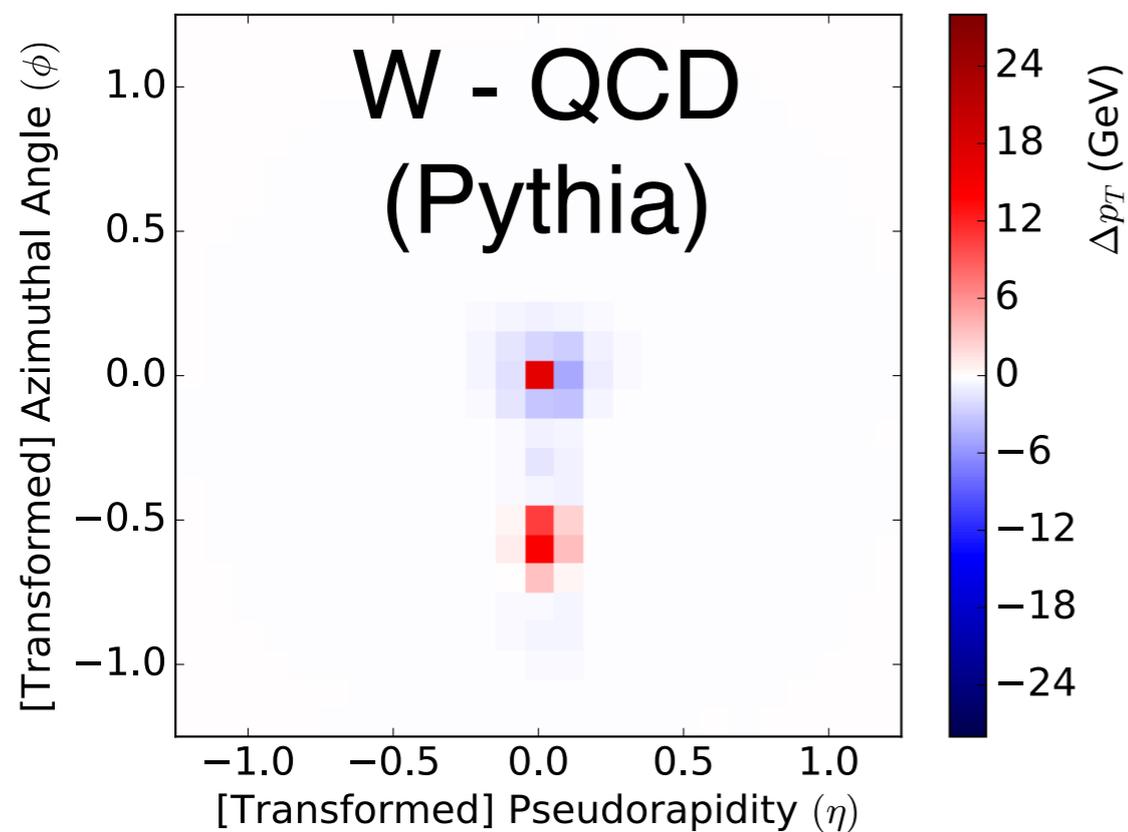
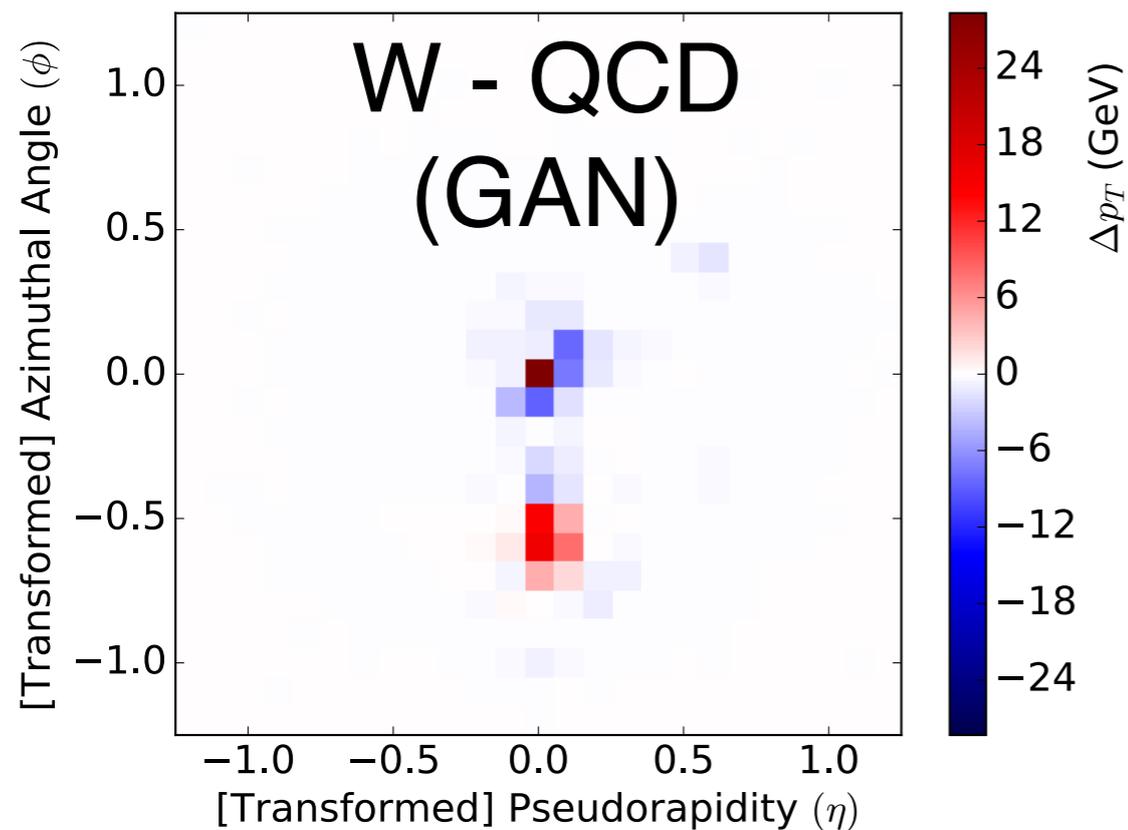
Jet images afford a lot of natural visualization



as a community, we have also developed many techniques



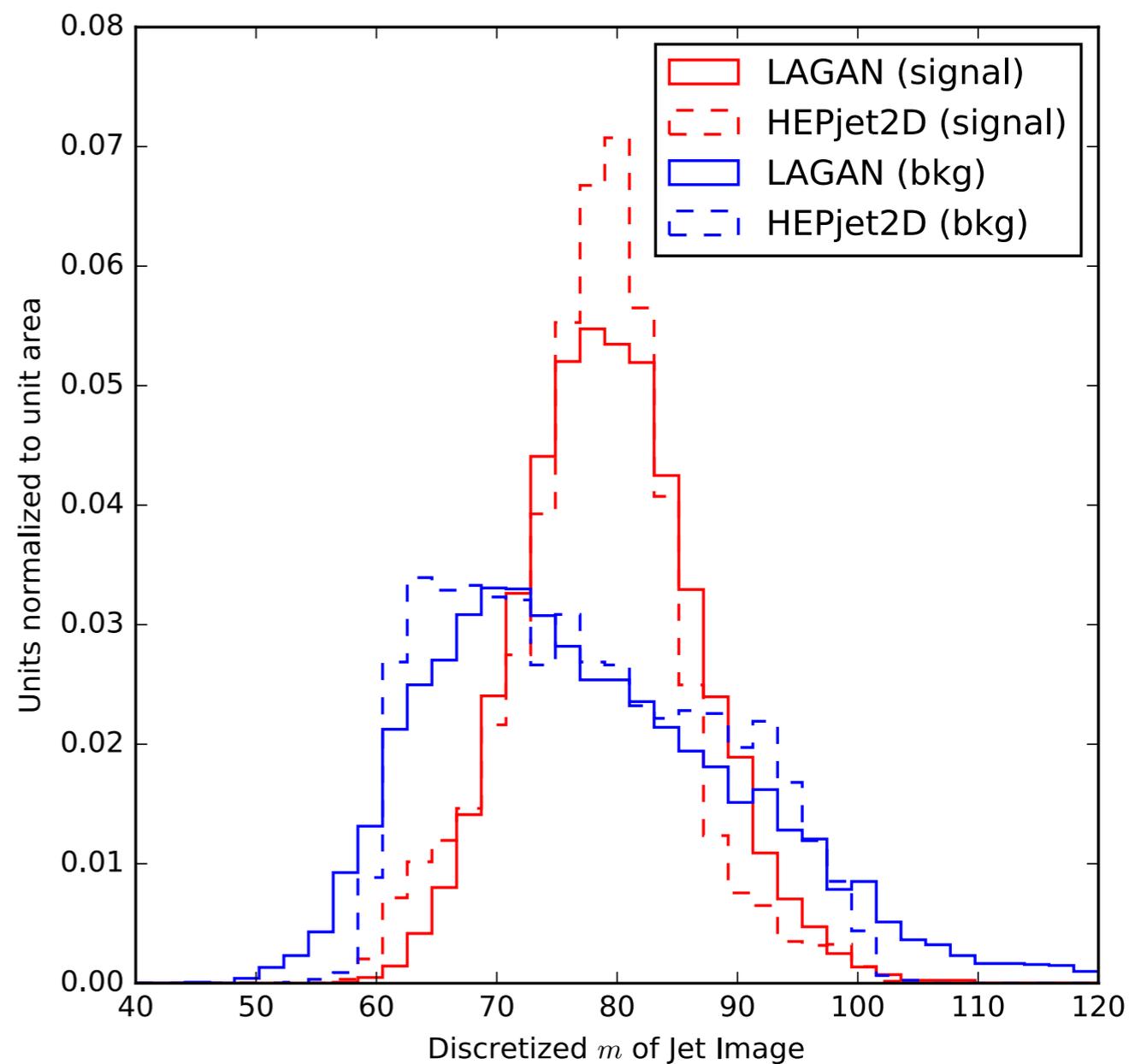
More detail in my [DS@HEP15](#) talk

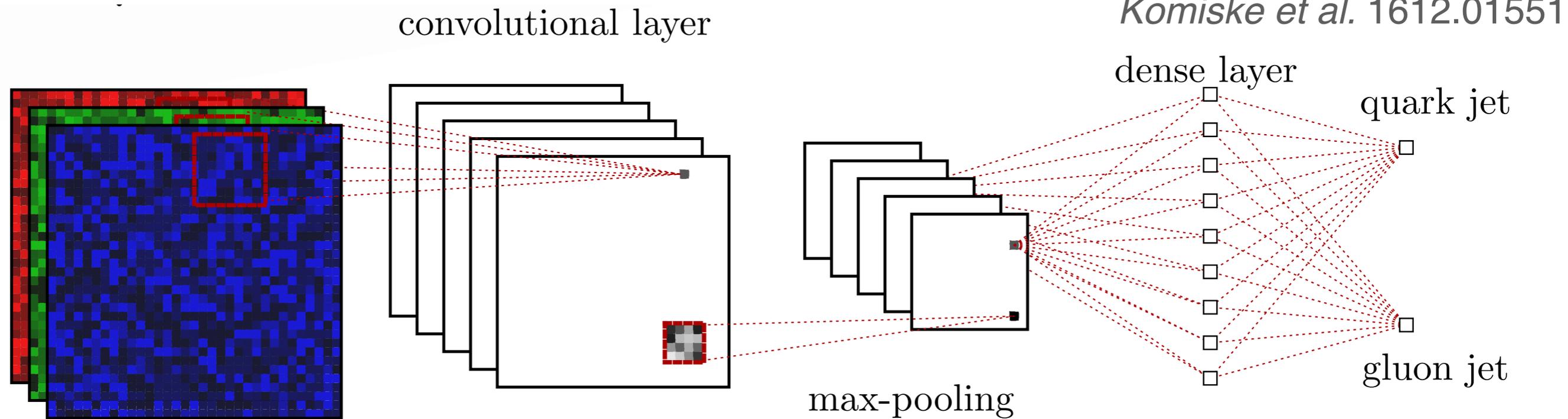


de Oliviera et al. arXiv:1701.05927

See [Michela's talk](#)

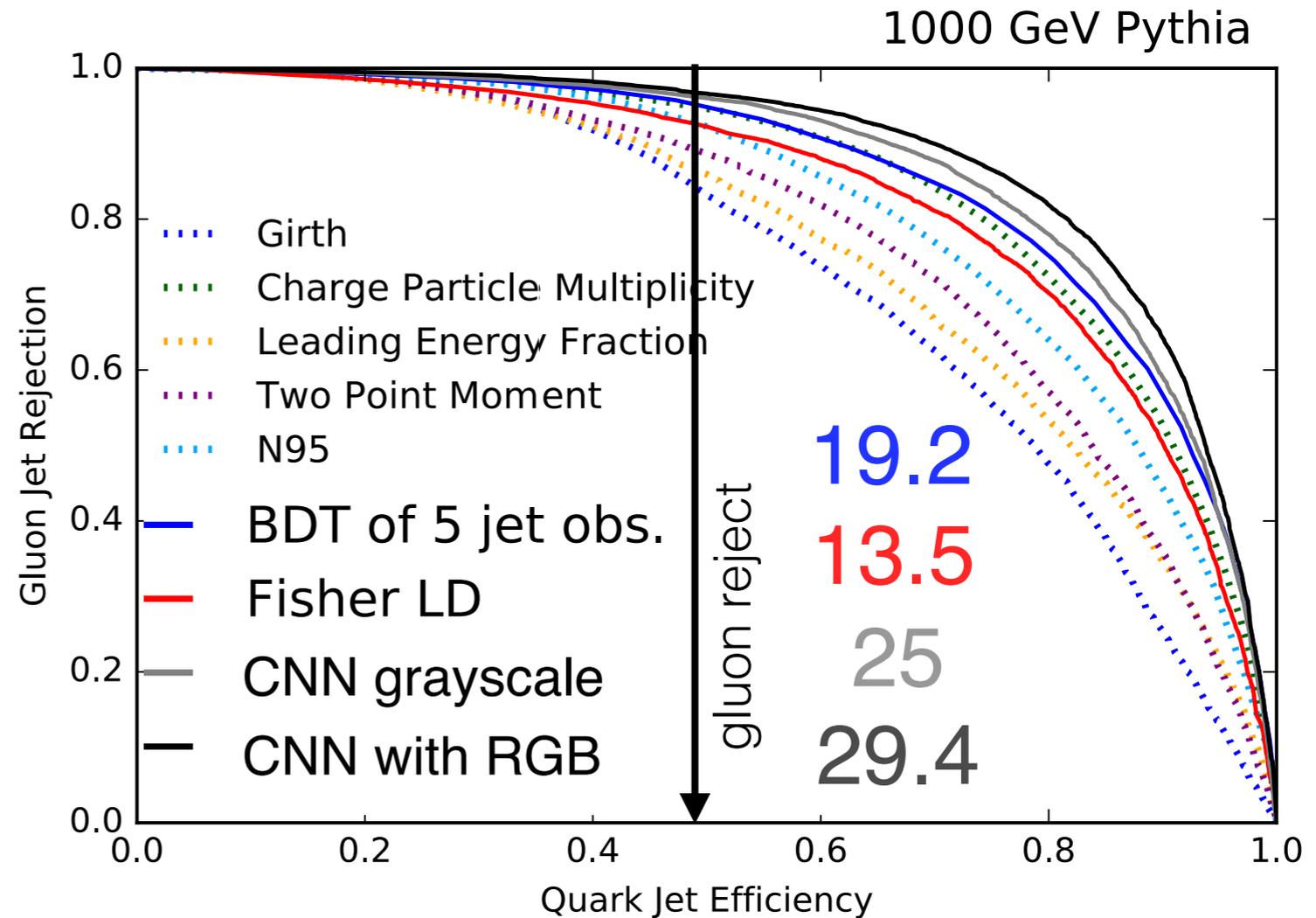
Generative Adversarial Networks (GANs)





red = charged p_T
green = neutral p_T
blue = n_{charged}

Fixed granularity for the three layers: **linear** < **shallow** < **deep** < **deep**

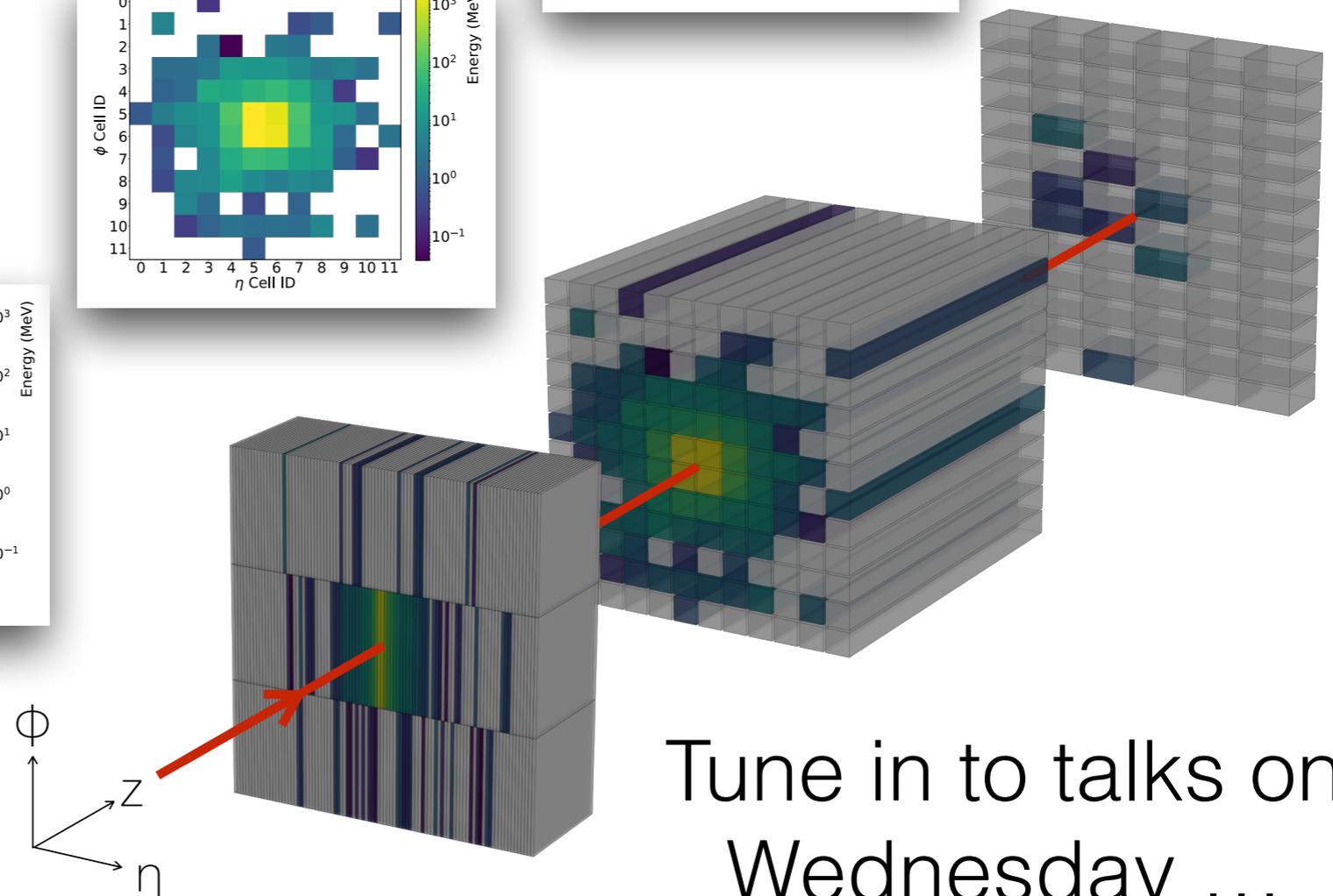
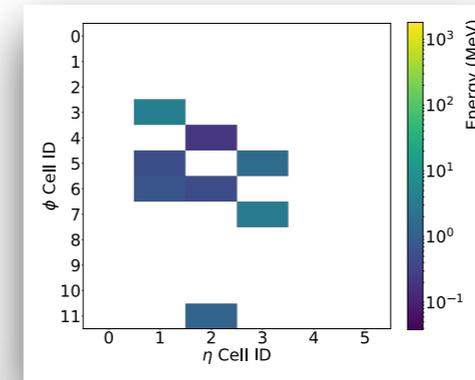
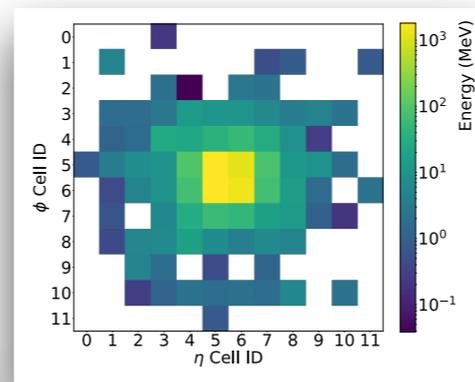
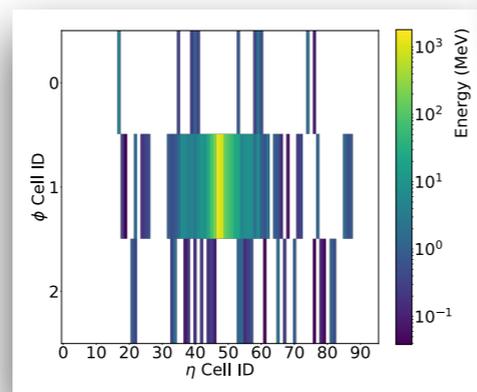


Paganini et al. arXiv:1705.tomorrow

See talks by Michela and Sofia

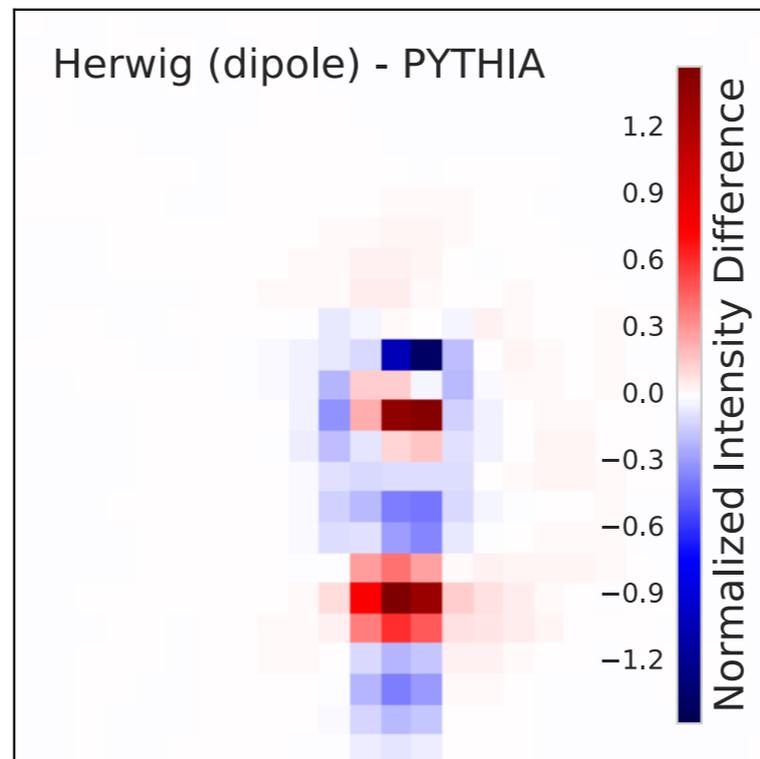
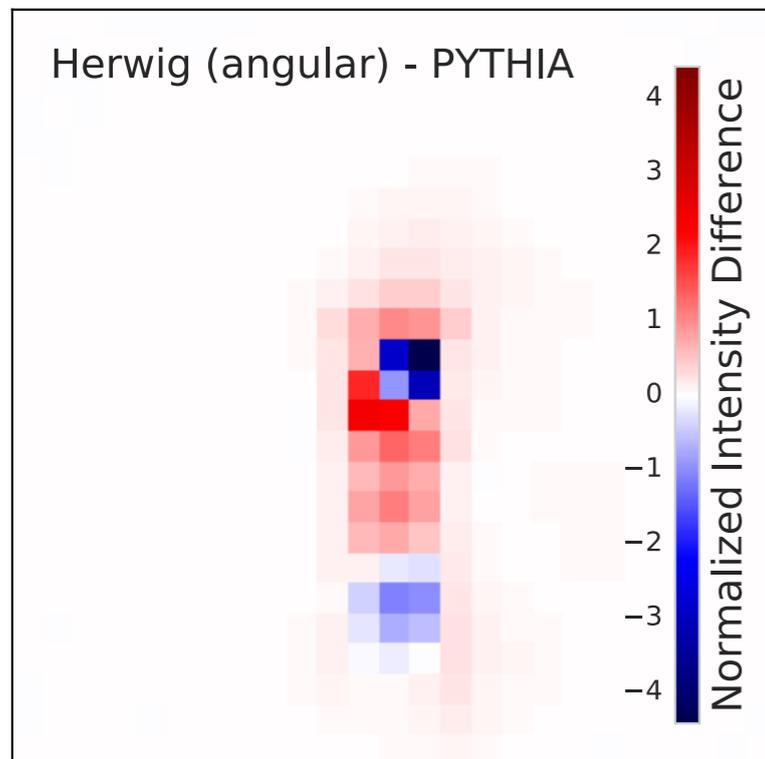
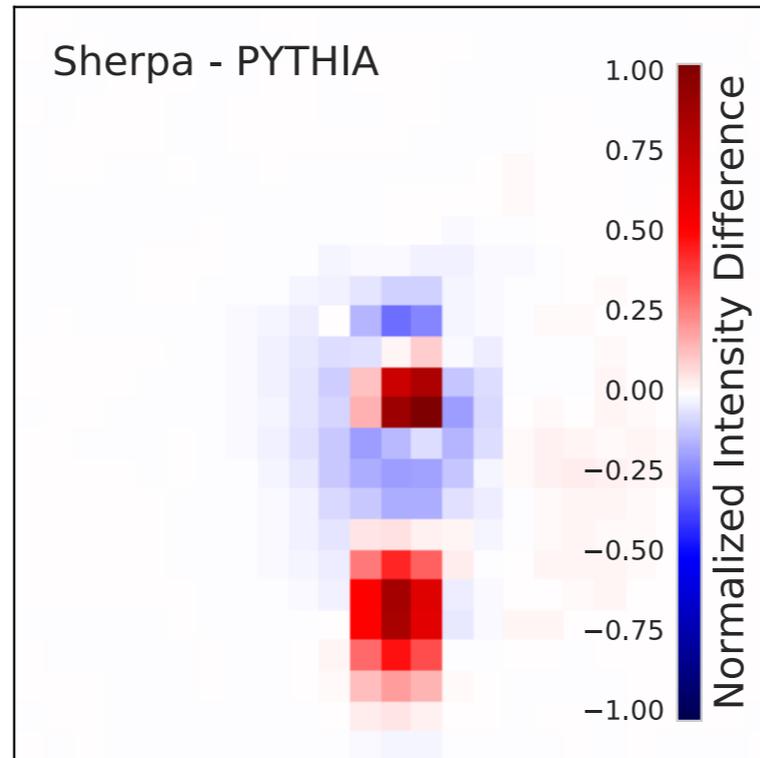
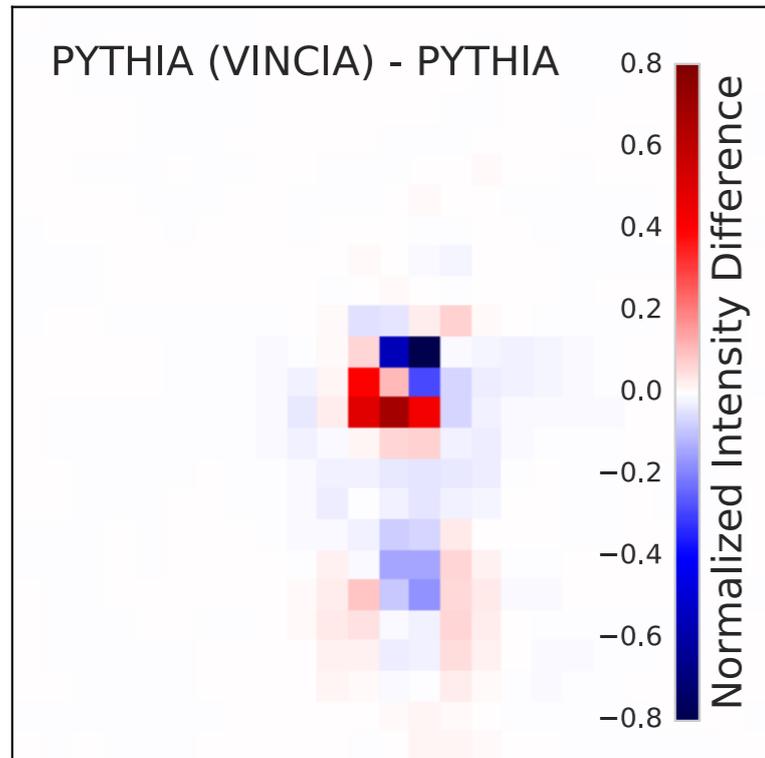
What about **multiple layers** with **non-uniform granularity** and a **causal relationship**?

Not jet images per se,
but the technology is
more general than jets!



Tune in to talks on
Wednesday ...

Boosted W boson jets



J. Barnard et al. 1609.00607

DNN classifiers
can **exploit**
subtle features

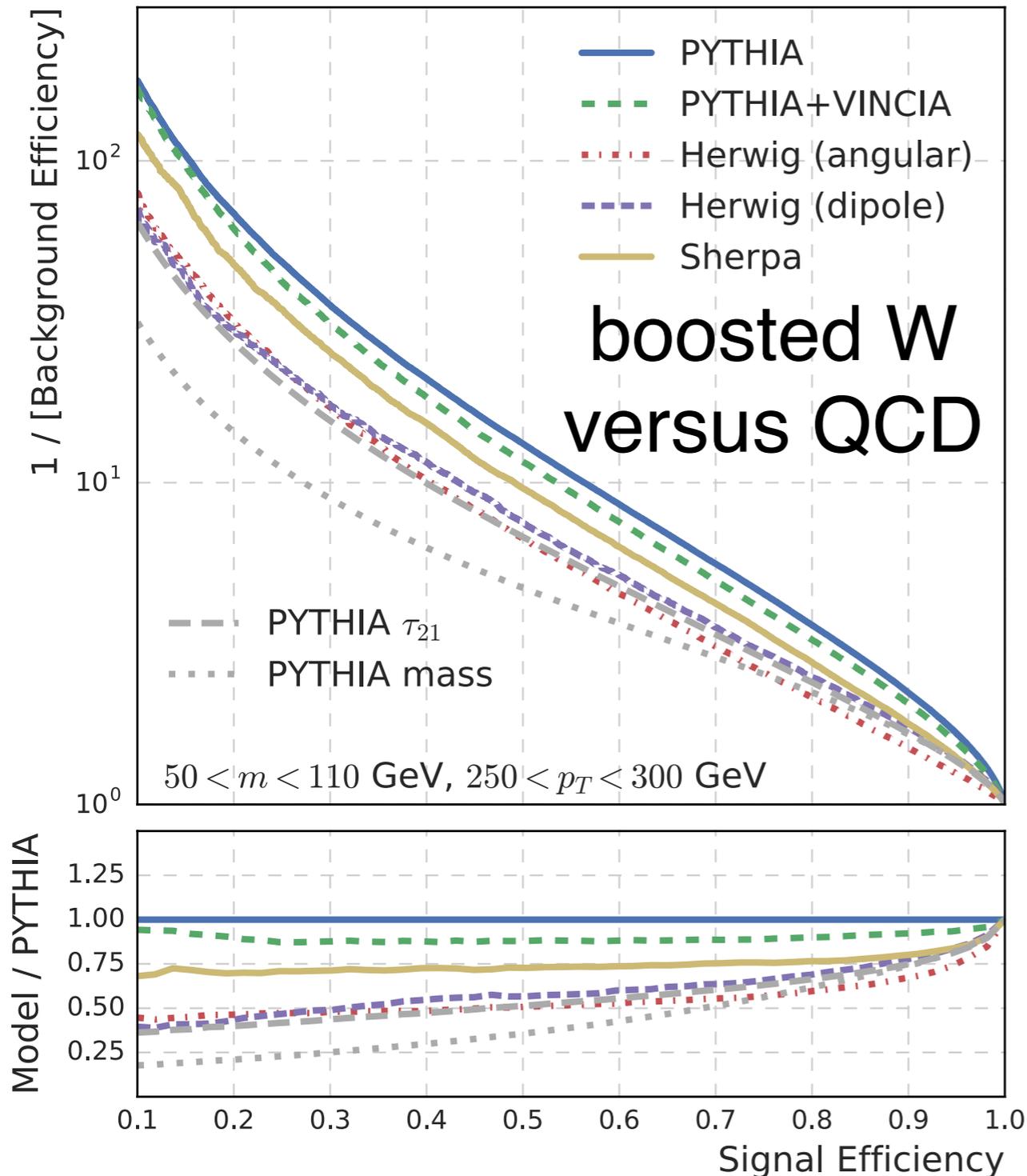
subtle features are
hard to model !

we need to be
careful about which
models we use -
only data is correct

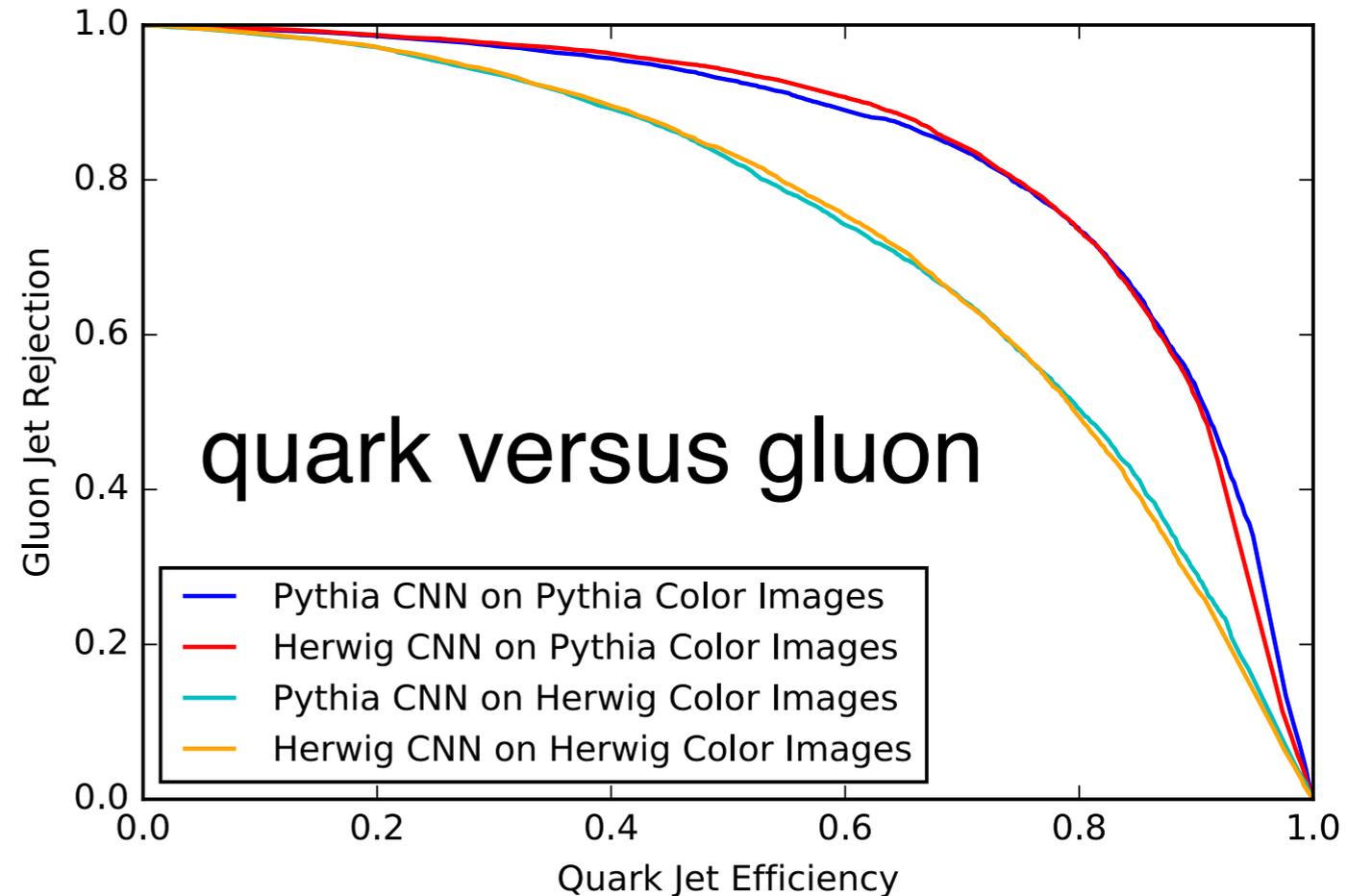
N.B. not all of these have been tuned to the same data

J. Barnard et al. 1609.00607

Deep Neural Network Performance



Komiske et al. 1612.01551



The story may be more complicated: even if performance differs, network itself may be similar

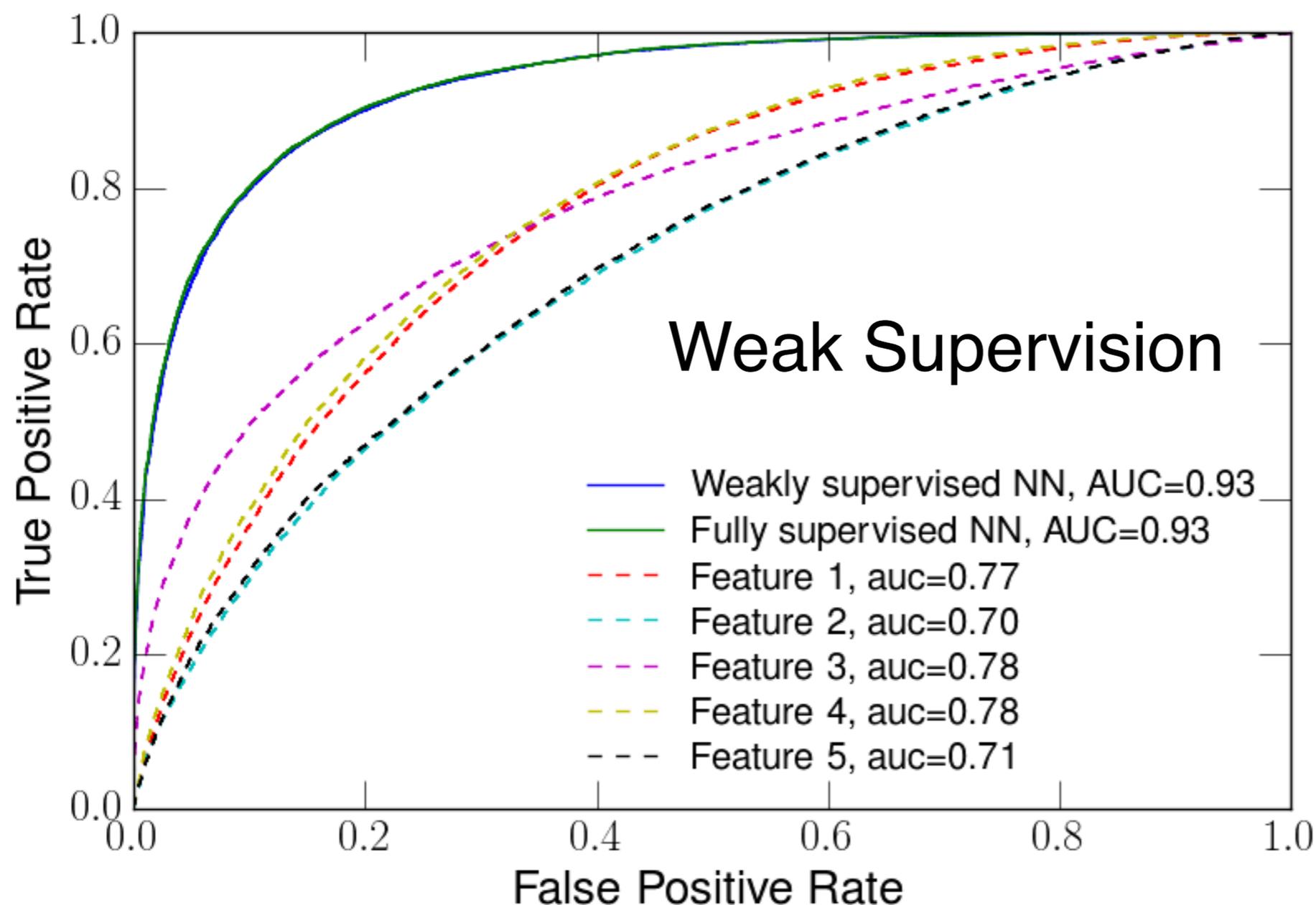
(could also be coincidence - fixed cuts may yield different performance)

The only way for your classifier to be **optimal when applied to data** is to **train on data**.

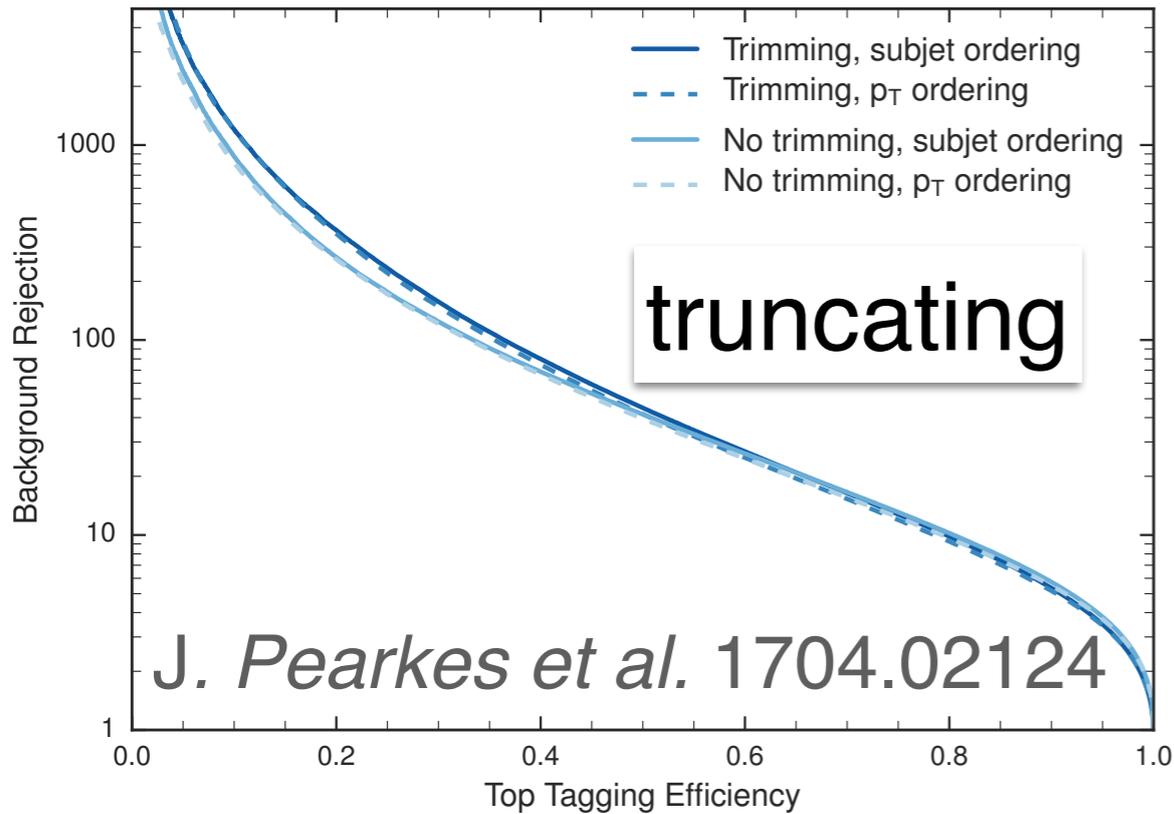
L. Dery et al. arXiv:1702.00414
see [Francesco's talk](#)

One regime where this is possible is when you have multiple samples with known class proportions.

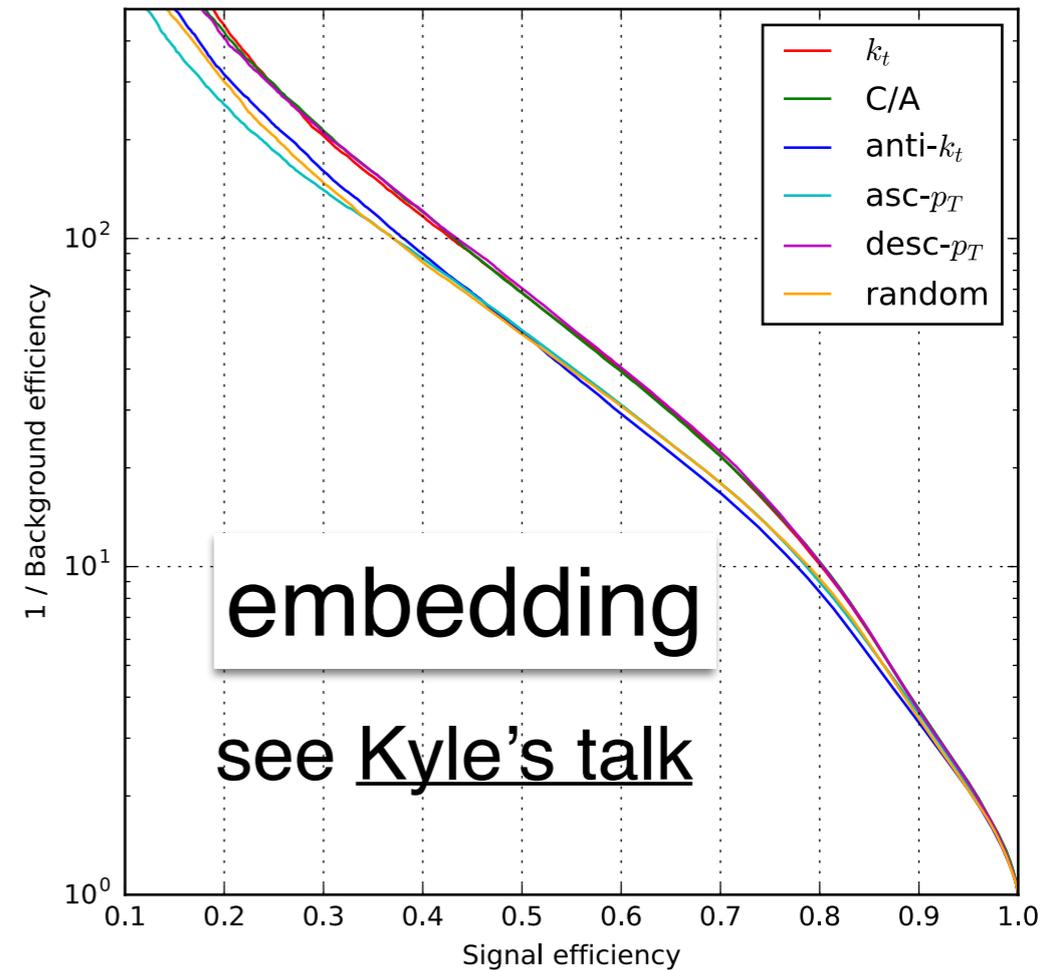
When the proportions are non-unity it is still possible to modify the loss and learn!



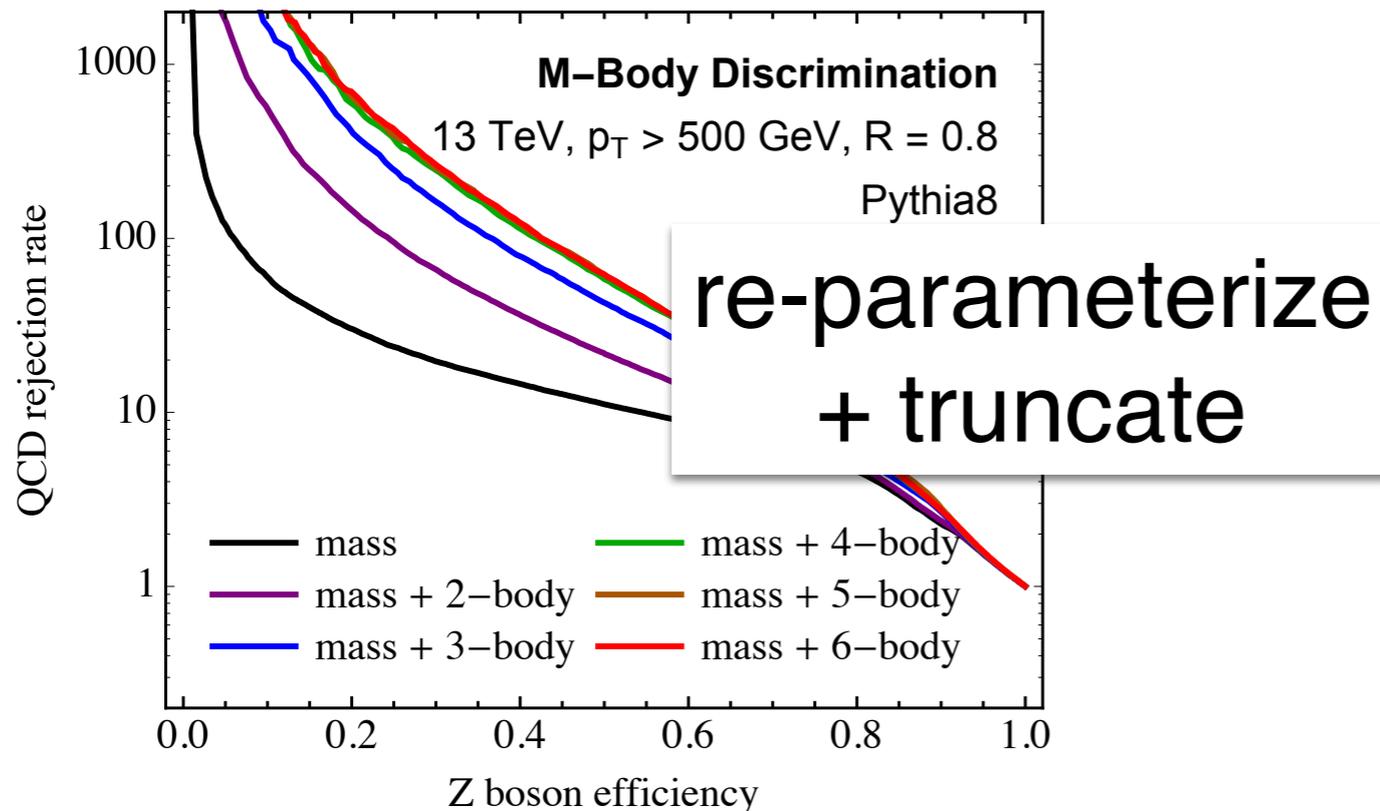
Jet $p_T = 600 - 2500$ GeV



G. Louppe et al. 1702.00748

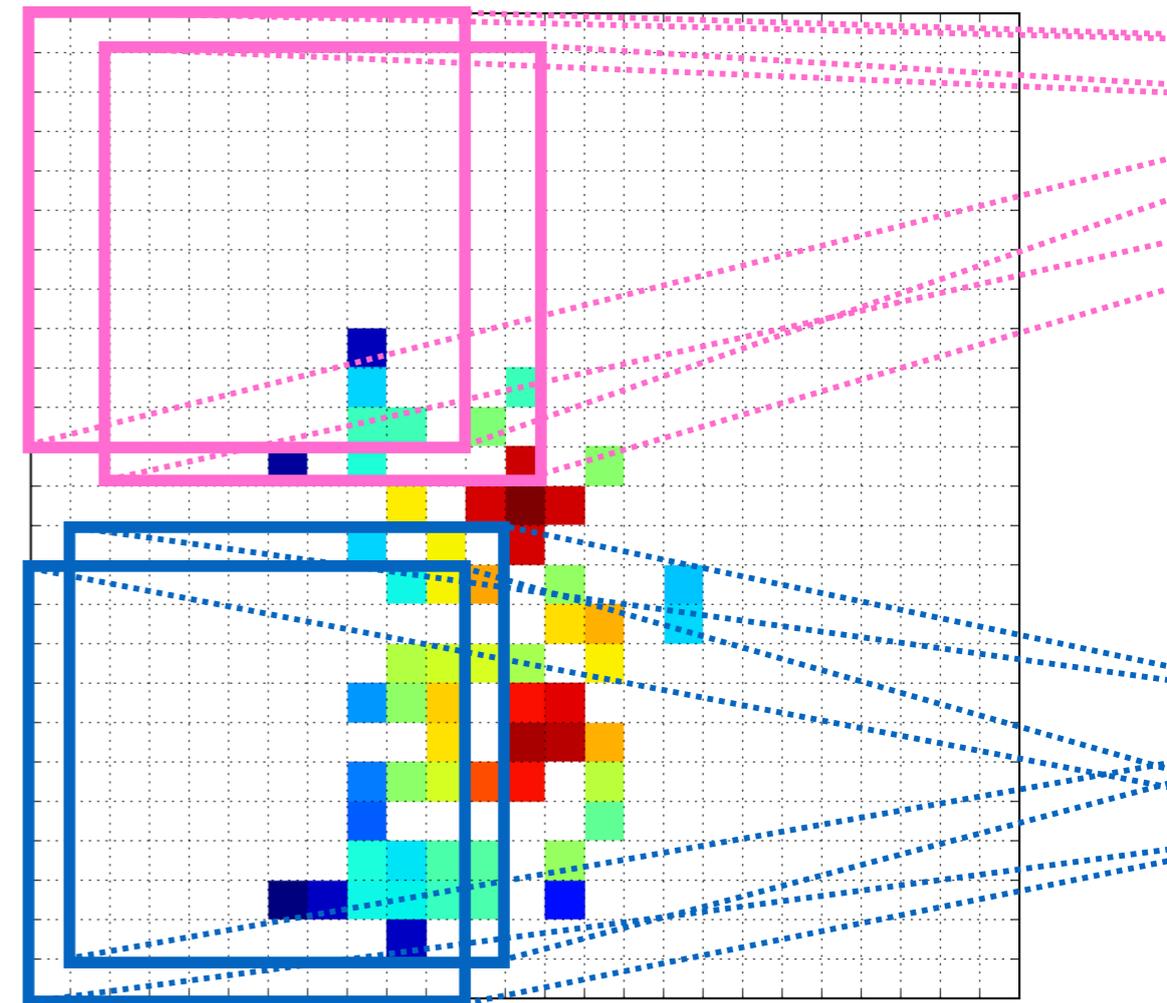
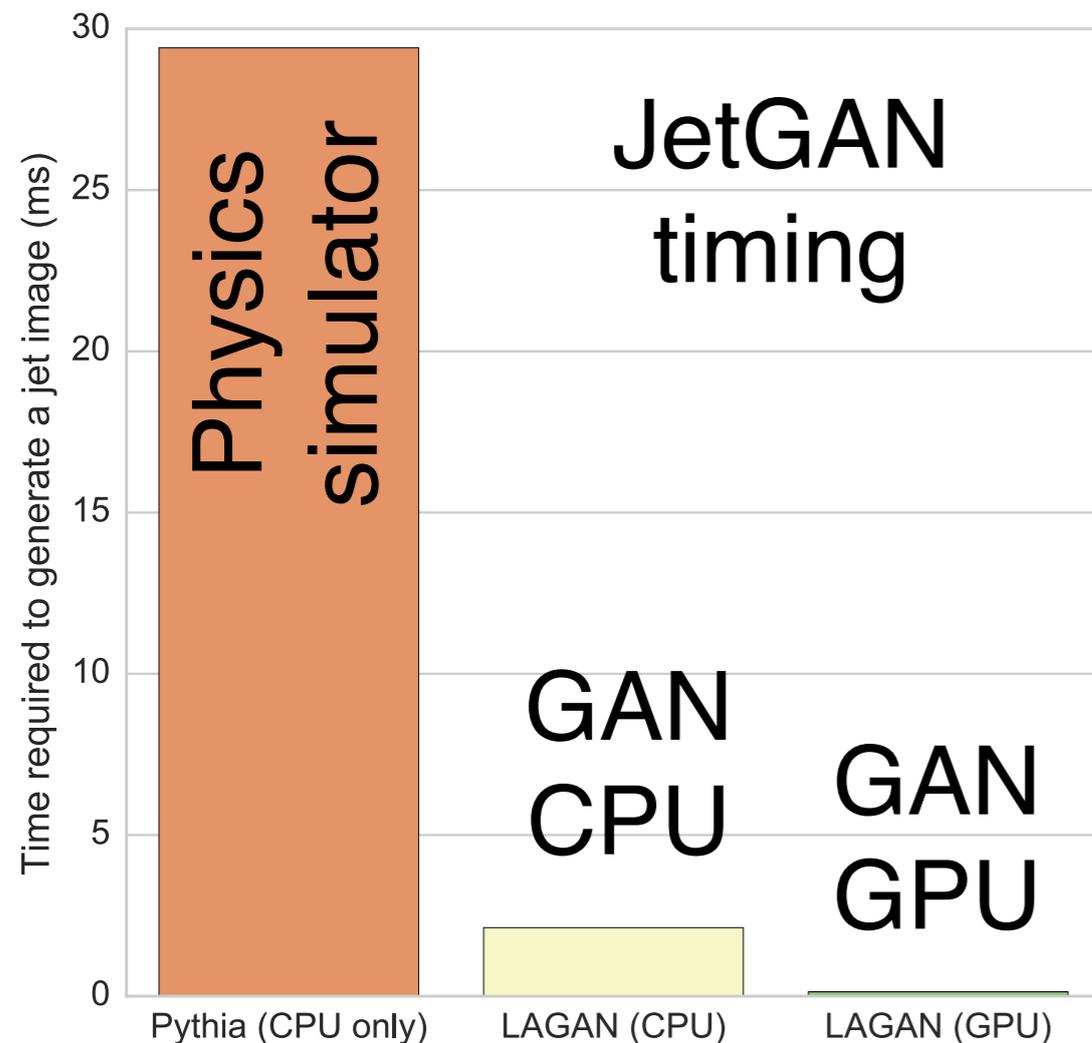


K. Datta et al. 1704.08249



Unfortunately, I don't have time to talk about these interesting approaches!

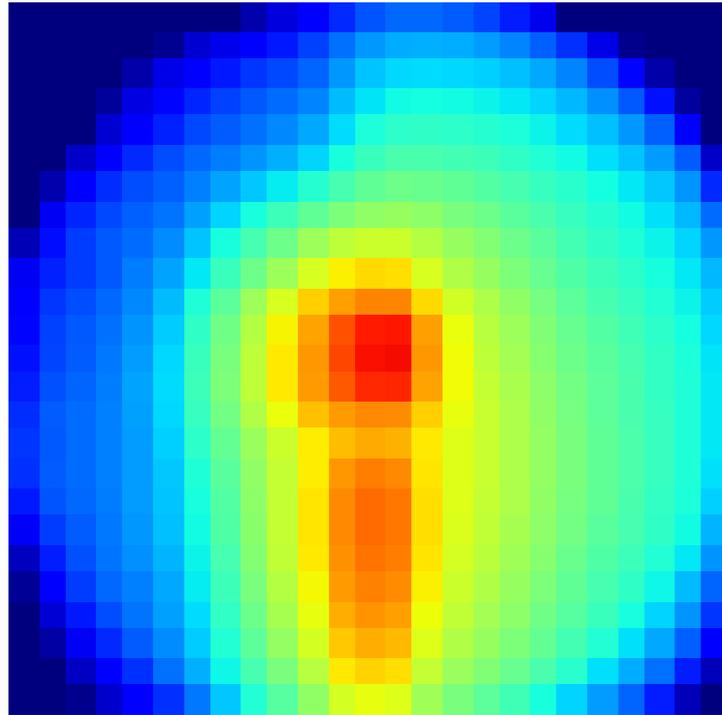
Jet image-based NN classification and generation are powerful tools for fully exploiting the physics program at the LHC



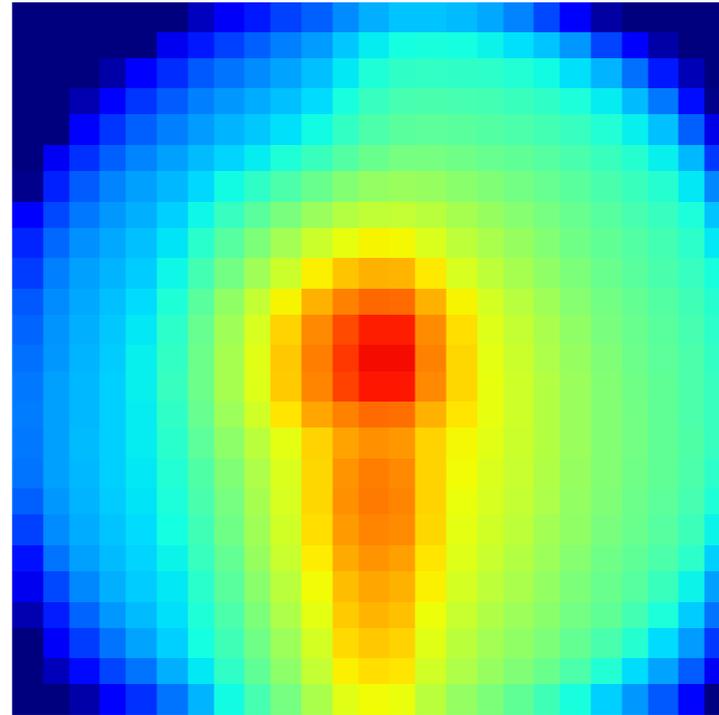
The key to robustness is to study what is being learned; this may even help us to learn something new about QCD!

Backup

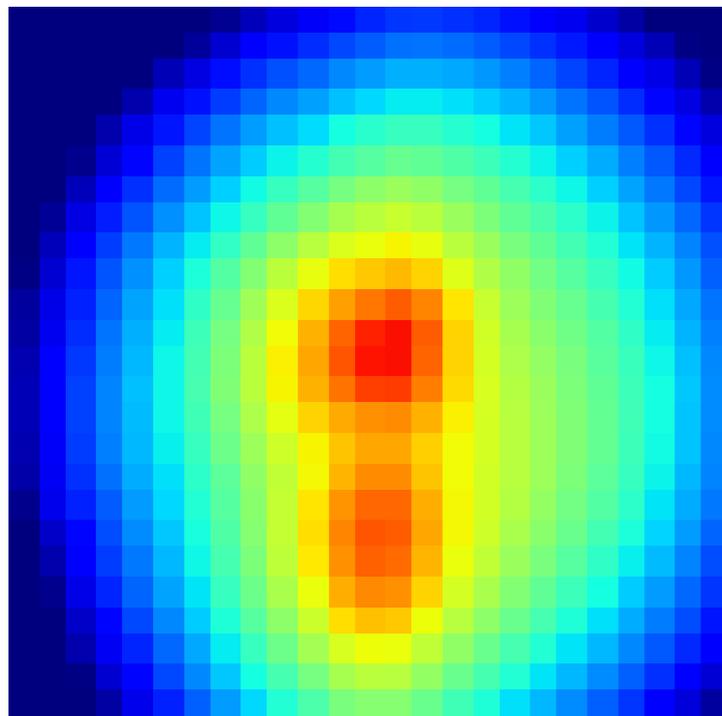
W



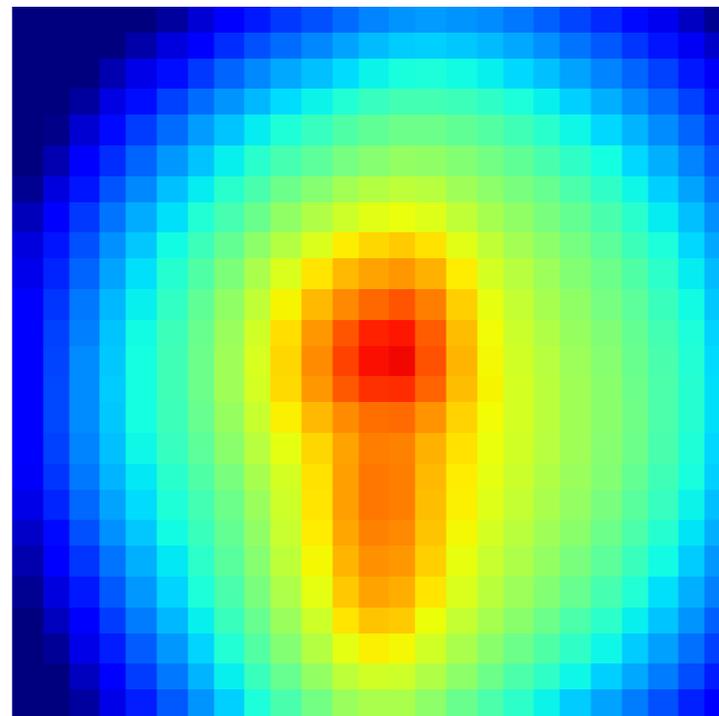
QCD



W Zoomed



QCD Zoomed

 $50 < m < 110 \text{ GeV}, 200 < p_T < 500 \text{ GeV}$

“Zooming” (two-prong)
J. Barnard et al. 1609.00607